

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**



REC'D 14 JUL 2000

WIPO

PCT

ESU

# BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

## COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 17 MARS 2000

Pour le Directeur général de l'Institut  
national de la propriété industrielle  
Le Chef du Département des brevets

DOCUMENT DE PRIORITÉ

PRÉSENTÉ OU TRANSMIS  
CONFORMÉMENT À LA  
RÈGLE 17.1.a) OU b)

Martine PLANCHE

INSTITUT  
NATIONAL DE  
LA PROPRIÉTÉ  
INDUSTRIELLE

SIEGE  
26 bis, rue de Saint Petersburg  
75800 PARIS Cedex 08  
Téléphone 01 53 04 53 04  
Télécopie 01 42 93 59 30



26 bis, rue de Saint Pétersbourg  
75800 Paris Cedex 08  
Téléphone : 01 53 04 53 04 Télécopie : 01 42 93 59 30

# BREVET D'INVENTION, CERTIFICAT D'UTILITE

Code de la propriété intellectuelle-Livre VI

cerja  
N° 55-1328

## REQUÊTE EN DÉLIVRANCE

Confirmation d'un dépôt par télécopie ☐

Cet imprimé est à remplir à l'encre noire en lettres capitales

Reserve à l'INPI

DATE DE REMISE DES PIÈCES **9 JUIN 1999**  
N° D'ENREGISTREMENT NATIONAL **9907250**  
DÉPARTEMENT DE DÉPÔT **75 INPI PARIS**  
DATE DE DÉPÔT **09 JUIN 1999**

1 NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE  
À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE  
**Hoechst Marion Roussel**  
**Monsieur VIEILLEFOSSE Jean Claude**  
**102, Route de Noisy**  
**93235 ROMAINVILLE CEDEX**

2 DEMANDE Nature du titre de propriété industrielle

☒ brevet d'invention ☐ demande divisionnaire  
☐ certificat d'utilité ☐ transformation d'une demande de brevet européen

☐ demande initiale  
☐ brevet d'invention

n° du pouvoir permanent **PG 06335** références du correspondant **ML/2517** téléphone **0149915727**

Établissement du rapport de recherche

☐ différé ☒ immédiat

Le demandeur, personne physique, requiert le paiement échelonné de la redevance

☐ oui ☒ non

Titre de l'invention (200 caractères maximum) **Nouveaux gènes de Candida albicans et les protéines codées par ces gènes.**

3 DEMANDEUR (S) n° SIREN **5 52 0 8 1 4 7 3**

code APE-NAP

Nom et prénoms (souligner le nom patronymique) ou dénomination

**Hoechst Marion Roussel**

Forme juridique

**Société Anonyme à**  
**Directoire et Conseil**  
**de Surveillance**

Nationalité (s) **FRANCAISE**

Adresse (s) complète (s)

**1, Terrasse Bellini**  
**92800 PUTEAUX**

Pays

**FRANCE**

En cas d'insuffisance de place, poursuivre sur papier libre ☐

4 INVENTEUR (S) Les inventeurs sont les demandeurs

☐ oui ☒ non Si la réponse est non, fournir une désignation séparée

5 RÉDUCTION DU TAUX DES REDEVANCES

☐ requise pour la 1ère fois ☐ requise antérieurement au dépôt ; joindre copie de la décision d'admission

6 DÉCLARATION DE PRIORITÉ OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE

pays d'origine

numéro

date de dépôt

nature de la demande

7 DIVISIONS

antérieures à la présente demande n°

date

n°

date

8 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE

(nom et qualité du signataire)

**Jean Claude VIEILLEFOSSE**

SIGNATURE DU PRÉPOSÉ À LA RECEPTION | SIGNATURE APRÈS ENREGISTREMENT DE LA DEMANDE À L'INPI



# BREVET D'INVENTION, CERTIFICAT D'UTILITE

## DÉSIGNATION DE L'INVENTEUR

(si le demandeur n'est pas l'inventeur ou l'unique inventeur)

### DIVISION ADMINISTRATIVE DES BREVETS

26bis, rue de Saint-Petersbourg  
75800 Paris Cédex 08  
Tél. : 01 53 04 53 04 - Télécopie 01 42 93 59 30

N° D'ENREGISTREMENT NATIONAL

99 07250 Cas 2517

**TITRE DE L'INVENTION:** Nouveaux gènes de *Candida albicans* et les protéines codées par ces gènes.

**LE(S) SOUSSIGNÉ(S)** Jean Claude VIEILLEFOSSE

**DÉSIGNE(NT) EN TANT QU'INVENTEUR(S)** (indiquer nom, prénoms, adresse et souligner le nom patronymique) :

- LALANNE Jean-Louis

110, Avenue du Maréchal  
94120 FONTENAY SOUS BOIS

- ROCHER Corinne

3, Rue Elisa Lemonnier  
75012 PARIS

NOTA : A titre exceptionnel, le nom de l'inventeur peut être suivi de celui de la société à laquelle il appartient (société d'appartenance) lorsque celle-ci est différente de la société déposante ou titulaire.

Date et signature (s) du (des) demandeur (s) ou du mandataire

Romainville, le 1er décembre 1999

  
Jean Claude VIEILLEFOSSE

# DOCUMENT COMPORTANT DES MODIFICATIONS

PAGE(S) DE LA DESCRIPTION OU DES REVENDICATIONS OU PLANCHE(S) DE DESSIN			R.M.*	DATE DE LA CORRESPONDANCE	TAMPON DATEUR DU CORRECTEUR
Modifiée(s)	Supprimée(s)	Ajoutée(s)			
p 36, 38, 39			R	10/10/99	J P M - 04 OCT. 1999

Un changement apporté à la rédaction des revendications d'origine, sauf si celui-ci découle des dispositions de l'article R.612-36 du code de la Propriété Intellectuelle, est signalé par la mention « R.M. » (revendications modifiées).

Nouveaux gènes de *Candida albicans* et  
les protéines codées par ces gènes.

La présente invention concerne de nouveaux gènes de  
5 *Candida albicans* et les protéines codées par ces gènes ainsi  
que les polynucléotides (ARN, ADN) codant pour ces protéines  
ou pour les polypeptides analogues de ces protéines.

La présente invention concerne également le procédé de  
préparation de ces polypeptides et polynucléotides, leur  
10 utilisation pour l'étude de mycètes pathogènes et notamment  
de *Candida albicans* et pour la préparation d'inhibiteurs des  
protéines codées par les gènes de la présente invention, ces  
inhibiteurs pouvant être utilisés comme agents antifongiques.  
La présente invention concerne également les compositions  
15 pharmaceutiques contenant de tels inhibiteurs.

La présente invention concerne donc notamment de  
nouvelles protéines de *Candida albicans* et les séquences  
nucléotidiques codant pour ces protéines, leur préparation et  
leurs utilisations.

20 Nous utiliserons également ci-après les abréviations  
suivantes : AA pour acides aminés, AN pour acides nucléiques,  
ARN pour acide ribonucléique, ARNm pour ARN messager, RNase  
pour ribonucléase, ADN pour acide désoxyribonucléique, ADNc  
pour ADN complémentaire, pb pour paires de bases, PCR pour  
25 réaction en chaîne par une polymérase, C.a. ou *C. albicans*  
pour *Candida albicans*, *E. coli* pour *Escherichia coli* et  
*S. cerevisiae* pour *Saccharomyces cerevisiae*.

Le terme criblage utilisé ci-après correspond au terme  
anglosaxon screening.

30 Le terme polynucléotides désigne ci-après les polynucléotides  
de la présente invention soit les séquences d'ADN et  
également ARN codant pour les protéines de la présente  
invention et leurs homologues codant pour des protéines de  
même fonction.

35 Le terme polypeptides désigne ci-après les polypeptides  
de la présente invention soit les protéines de la présente  
invention et leurs analogues ou homologues fonctionnels tels  
que définis ci-après, ayant donc les mêmes fonctions.

Le terme mycète désigne ci-après un organisme eucaryote, porteur de spores, dont la nutrition se fait par absorption, qui est dépourvu de chlorophylle et qui se reproduit de façon sexuée ou asexuée.

5 Les mycoses sont des infections de l'homme ou des animaux qui peuvent être superficielles ou profondes, causées par des champignons pathogènes. Dans le cas de mycoses profondes, elles peuvent être très sévères et de pronostic grave.

10 Des substances antimycotiques à effets fongistatiques ou fongicides sont utilisées dans le traitement des mycoses. Ce traitement est difficile car il existe peu de substances antifongiques disponibles pour la thérapeutique et qu'elles ont souvent des effets secondaires qui limitent leur  
15 utilisation. Par exemple, l'Amphotéricine B, qui représente le traitement de choix des mycoses profondes, a des effets secondaires néphrotoxiques.

Il existe donc une forte demande pour de nouvelles substances efficaces contre les champignons pathogènes et  
20 susceptibles d'être utilisées en thérapeutique contre les infections fongiques. Ces substances pourront être utilisées soit en prophylaxie, dans le cas des états d'immunodépression graves soit en traitement curatif des infections fongiques. De plus, ces substances devront avoir un mode d'action  
25 spécifique, leur permettant d'inhiber la croissance ou de tuer les cellules de mycètes sans altérer les fonctions essentielles des cellules humaines.

L'objet de la présente invention est de proposer des gènes pouvant constituer de nouvelles cibles pour  
30 l'identification de substances antifongiques et notamment de substances permettant de traiter les infections dues aux champignons du genre Candida.

Ces gènes seront notamment des gènes essentiels indispensables à la survie et à la multiplication des  
35 cellules.

Différentes méthodes peuvent être utilisées pour déterminer si le produit d'un gène est essentiel à la survie d'un mycète ou essentiel à l'établissement ou au maintien

d'une infection. L'identification du caractère essentiel d'un gène apporte une information additionnelle concernant sa fonction et permet de sélectionner les gènes dont le produit constitue une cible intéressante pour une substance

5 antifongique. Des exemples de ces méthodes sont résumés brièvement ci-après. Ces méthodes sont décrites dans les ouvrages suivants :

- Guthrie C. and Fink G.R. Eds. *Methods in Enzymology*, Vol 194, 1991, 'Guide to Yeast Genetics and Molecular Biology', Academic Press Inc.
- 10 - Rose A.H., A.E. Wheals and J.S. Harrison Eds. *The yeasts*, Vol.6, 1995, 'Yeast Genetics' , Academic Press Inc.
- Ausubel F. et al. Eds. 'Short Protocols in Molecular Biology', 1995, Wiley.
- 15 - Brown A.J.P. and Tuite M.F. (Eds) 'Yeast Gene Analysis' *Methods in Microbiology*, Vol 26, 1998, Academic Press Inc.

Selon les cas, on utilisera l'une ou l'autre des méthodes décrites en fonction du résultat recherché. Notamment, on pourra procéder par une méthode d'inactivation

20 directe du gène ou d'inactivation transitoire du gène.

Dans la levure *S. cerevisiae*, la méthode la plus couramment utilisée consiste à inactiver le gène étudié dans le chromosome de la levure. L'allèle sauvage est inactivé par insertion d'un marqueur génétique (par exemple un gène

25 d'auxotrophie ou un marqueur de résistance). Cette insertion est obtenue en général par la méthode de conversion génique à l'aide de cassettes de délétion linéaires préparées selon les méthodes connues telles que décrites dans Guthrie C. and Fink G.R. Eds. *Methods in Enzymology*, Vol 194, 1991, 'Guide to Yeast Genetics and Molecular Biology', Academic Press Inc. ou

30 dans Gultner et al. *Nucleic Acid Research*, 1996, 24: 2519-2524.

L'inactivation se fait dans une souche diploïde puis la méiose est induite par des méthodes classiques comme par

35 exemple la croissance en milieu pauvre en azote et les quatre spores issues d'asques individuels sont isolées par micromanipulation. L'inactivation d'un gène essentiel se traduit par une perte de viabilité des deux spores (sur



quatre) qui ont acquis le marqueur de sélection. La viabilité de ces spores peut être restaurée par l'introduction dans la souche d'un plasmide centromérique ou réplcatif portant une copie du gène sauvage.

- 5 On peut également procéder par inactivation transitoire du gène : l'utilisation de promoteurs régulables permet également de déterminer si un gène est essentiel à la survie d'une cellule. Pour ce faire, on remplace le promoteur natif du gène par un promoteur régulable directement sur le
- 10 chromosome ou sur un plasmide extra-chromosomique. On peut par exemple utiliser le promoteur GAL ou ses dérivés ou le promoteur tetO (Mumberg et al. 1994, Nucleic Acid Research, 22 : 5767-5768 ; Belli et al. 1998, Yeast, 14 : 1127-1138). Le caractère essentiel du gène étudié peut ainsi être observé
- 15 lorsque le promoteur utilisé est réprimé, soit dans les souches haploïdes chez la levure *S. cerevisiae*, soit après inactivation du deuxième allèle chez les micro-organismes diploïdes tels que *C. albicans*.

- A partir d'un gène essentiel connu dans une espèce, on
- 20 peut procéder à l'identification de gènes homologues ou de même fonction dans une autre espèce de mycète : les méthodes connues peuvent être utilisées pour identifier les gènes homologues d'un gène étudié dans une autre espèce de mycète (gènes dits 'orthologues') ou les gènes de même fonction que
- 25 le gène étudié. Des exemples de méthodes utilisables sont développées ci-après. Ces méthodes sont décrites dans les ouvrages suivants :

Sambrook et al. 1989, Molecular Cloning, Cold Spring Harbor Laboratory Press.

- 30 - Ausubel F. et al. Eds. 'Short Protocols in Molecular Biology', 1995, Wiley.

- Guthrie C. and Fink G.R. Eds. Methods in Enzymology, Vol 194, 1991, 'Guide to Yeast Genetics and Molecular Biology', Academic Press Inc.

- 35 On peut procéder par exemple par criblage par homologie, par complémentation génique ou encore par amplification par PCR en utilisant des amorces spécifiques à partir de banques d'ADN génomique ou de banques d'ADN complémentaire (ADNc) des

mycètes pathogènes.

Les banques d'ADN génomique ou d'ADNc peuvent être préparées selon les méthodes connues et les fragments polynucléotidiques obtenus sont intégrés dans un vecteur d'expression, par exemple un vecteur tel que pRS423 ou ses dérivés qui sont utilisables aussi bien dans la bactérie *E. coli* que dans *S. cerevisiae*. Le criblage de la banque se fera par les méthodes classiques d'hybridation *in situ* sur une réplique des colonies bactériennes. Les conditions d'hybridation seront adaptées à la stringence voulue pour la réaction, de façon à identifier des fragments de plus ou moins grande homologie avec le gène étudié.

Les gènes des autres espèces de mycètes peuvent également être identifiés par des méthodes connues dites de 'complémentation génique'. Par exemple, une souche de *S. cerevisiae* dans laquelle un gène essentiel identifié a été placé sous le contrôle d'un promoteur régulable peut être transformée par un échantillon représentatif d'une banque d'ADN ou d'ADNc correspondant au mycète étudié tel que *C. albicans*. Lorsque les levures sont cultivées dans des conditions telles que le promoteur est réprimé, seules peuvent survivre les levures portant un vecteur recombinant contenant une séquence du mycète étudié fonctionnellement équivalente au gène essentiel initial. La séquence du gène dans le mycète étudié est ensuite identifiée en isolant le vecteur recombinant et en le séquençant selon les méthodes connues. De la même façon, la méthode dite de 'plasmid shuffle' permet de sélectionner les levures ayant perdu l'expression du gène essentiel initial et contenant une séquence fonctionnellement équivalente provenant d'un autre mycète.

L'étude peut être réalisée sur différentes espèces : les gènes fonctionnellement équivalents ou homologues en séquence à un gène essentiel peuvent être isolés dans d'autres mycètes et notamment dans les différents mycètes pathogènes pour l'homme. Pour cela peuvent être utilisés notamment les mycètes appartenant aux classes Zygomycètes, Basidiomycètes, Ascomycètes et Deutéromycètes. Tout particulièrement, les

mycètes appartiendront aux sous-classes *Candida* spp.,  
notamment *Candida albicans*, *Candida glabrata*, *Candida*  
*tropicalis*, *Candida parapsilosis* et *Candida krusei*. Les  
mycètes appartiendront également aux sous-classes *Aspergillus*  
5 *fumigatus*, *Coccidioides immitis*, *Cryptococcus neoformans*,  
*Histoplasma capsulatum*, *Blastomycès dermatidis*,  
*Paracoccidioides brasiliensis* et *Sporothrix schenckii*.

La présente invention concerne ainsi l'identification de  
substances antimycotiques telles que notamment de substances  
10 anti-*Candida albicans*.

La présente invention concerne ainsi des inhibiteurs de  
protéines fongiques pouvant être utilisés comme agents  
antifongiques.

On connaît ainsi des organismes pathogènes tels que la  
15 levure pathogène *Candida albicans* qui causent des maladies  
infectieuses dans l'organisme humain. Dans le but de trouver  
des moyens de traiter des maladies, on peut choisir des  
cibles telles que par exemple intracellulaires et l'une ou  
plusieurs des protéines de la présente invention codées par  
20 les gènes de la présente invention peut ou peuvent être l'une  
ou certaines de ces cibles.

La présente invention a ainsi permis d'isoler des  
polynucléotides ADN et ARN codant pour des protéines de  
*Candida albicans* et de révéler leurs séquences  
25 nucléotidiques.

Nous appellerons les gènes de la présente invention  
codant pour les protéines de *Candida albicans* de la présente  
invention comme suit : CaDR472, CaDR489, CaDR527 sous forme  
de deux allèles différents soit 1CaDR527 et 2CaDR527,  
30 CaFL024, CaNL260 et CaDR361.

Les séquences nucléotidiques de ces gènes (et des deux  
allèles pour CaDR527) sont donnés dans le listing de  
séquences ci-après et sont respectivement nommés comme suit :

- SEQ ID N° 1 pour CaDR472,
- 35 - SEQ ID N° 3 pour CaDR489,
- SEQ ID N° 5 pour le 1er allèle de CaDR527 soit 1CaDR527,
- SEQ ID N° 7 pour le 2ème allèle de CaDR527 soit 2CaDR527,
- SEQ ID N° 9 pour CaFL024,

- SEQ ID N° 11 pour CaNL260
- et SEQ ID N° 13 pour CaDR361.

Les séquences polypeptidiques des protéines codées par les gènes de la présente invention sont respectivement

5 nommées comme suit :

- SEQ ID N° 2 ou PCaDR472 pour la protéine codée par CaDR472,
- SEQ ID N° 4 ou PCaDR489 pour la protéine codée par CaDR489,
- 10 - SEQ ID N° 6 ou 1PCaDR527 pour la protéine codée par 1CaDR527,
- SEQ ID N° 8 ou 2PCaDR527 pour la protéine codée par 2CaDR527,
- SEQ ID N° 10 ou PCaFL024 pour la protéine codée par CaFL024,
- 15 - SEQ ID N° 12 ou PCaNL260 pour la protéine codée par CaNL260
- et SEQ ID N° 14 ou PCaDR361 pour la protéine codée par CaDR361.

20 La présente invention a donc pour objet des polynucléotides isolés contenant chacun une séquence nucléotidique choisie dans le groupe suivant :

- a) un polynucléotide ayant au moins 50 % ou au moins 60 % et de préférence au moins 70 % d'identité avec un
- 25 polynucléotide codant pour un polypeptide ayant la même fonction et ayant une séquence en acides aminés homologue d'une séquence choisie parmi SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14, telles que définies ci-dessus et ci-après,
- 30 b) un polynucléotide complémentaire du polynucléotide a)
- c) un polynucléotide comprenant au moins 15 bases consécutives du polynucléotide défini en a) et b).

La présente invention a ainsi pour objet les polynucléotides définis ci-dessus tels que ces

35 polynucléotides sont des ADN.

La présente invention a ainsi pour objet les polynucléotides définis ci-dessus tels que ces polynucléotides sont des ARN.

La présente invention a plus précisément pour objet les polynucléotides tels que définis ci-dessus comprenant chacun une séquence de nucléotides choisie parmi SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 et  
5 SEQ ID N° 13 telles que définies ci-dessus et ci-après.

La présente invention a ainsi permis d'isoler les séquences d'ADN codant respectivement pour les protéines de *Candida albicans* PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361, telles que définies ci-dessus.

10 La présente invention a également permis de révéler les séquences d'acides nucléiques des gènes de la présente invention et également les séquences d'acides aminés des protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361, codées par ces gènes.

15 La présente invention a ainsi pour objet les séquences d'ADN telles que définies par les polynucléotides ci-dessus, caractérisées en ce que ces séquences d'ADN sont celles des gènes codant respectivement pour des protéines de *Candida albicans* (ayant les mêmes fonctions que les protéines  
20 PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361) et contenant chacune une séquence de nucléotides choisie parmi SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 et SEQ ID N° 13 telles que définies ci-dessus et ci-après.

25 Une telle séquence SEQ ID N° 1 de la présente invention comprend donc 747 nucléotides.

Une telle séquence SEQ ID N° 3 de la présente invention comprend donc 711 nucléotides.

30 Une telle séquence SEQ ID N° 5 de la présente invention comprend donc 1383 nucléotides.

Une telle séquence SEQ ID N° 7 de la présente invention comprend donc 1383 nucléotides.

Une telle séquence SEQ ID N° 9 de la présente invention comprend donc 2262 nucléotides.

35 Une telle séquence SEQ ID N° 11 de la présente invention comprend donc 447 nucléotides.

Une telle séquence SEQ ID N° 13 de la présente invention comprend donc 966 nucléotides.

La présente invention a aussi pour objet les séquences d'ADN de gènes telles que définies ci-dessus codant chacune pour une séquence d'acides aminés choisie parmi SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14.

La séquence SEQ ID N° 2 de la protéine PCaDR472 comprend donc 248 AA.

La séquence SEQ ID N° 4 de la protéine PCaDR489 comprend donc 236 AA.

10 La séquence SEQ ID N° 6 de la protéine 1PCaDR527 comprend donc 460 AA.

La séquence SEQ ID N° 8 de la protéine 2PCaDR527 comprend donc 460 AA.

La séquence SEQ ID N° 10 de la protéine PCaFL024  
15 comprend donc 753 AA.

La séquence SEQ ID N° 12 de la protéine PCaNL260 comprend donc 148 AA.

La séquence SEQ ID N° 14 de la protéine PCaDR361 comprend donc 321 AA.

20 La présente invention a particulièrement pour objet les séquences d'ADN codant pour les protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 telles que définies ci-dessus ainsi que les séquences d'ADN qui hybrident avec celles-ci et/ou présentent des homologies  
25 significatives avec ces séquences ou des fragments de celles-ci et codent pour des protéines ayant les mêmes fonctions.

La présente invention a également pour objet les séquences d'ADN telles que définies ci-dessus comprenant des modifications introduites par suppression, insertion et/ou  
30 substitution d'au moins un nucléotide codant pour des protéines ayant les mêmes activités que les protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 telles que définies ci-dessus.

La présente invention a notamment pour objet les  
35 séquences d'ADN telles que définies ci-dessus ainsi que les séquences d'ADN qui ont une homologie de séquence nucléotidique d'au moins 50 % ou au moins 60 % et de préférence au moins 70 % avec lesdites séquences d'ADN.

La présente invention a ainsi également pour objet les séquences d'ADN telles que définies ci-dessus ainsi que les séquences d'ADN qui codent pour des protéines de fonctions similaires dont les séquences respectives en AA ont une  
 5 homologie d'au moins 40 % et notamment de 45 % ou d'au moins 50 %, plutôt au moins 60 % et de préférence au moins 70 % avec les séquences en AA codées par lesdites séquences d'ADN.

Par séquences qui hybrident, on inclut les séquences d'ADN qui hybrident avec l'une des séquences d'ADN ci-dessus  
 10 sous des conditions standard de stringence élevée, moyenne ou basse et qui codent pour un polypeptide ayant la même fonction. Les conditions de stringence sont celles réalisées dans les conditions connues de l'homme du métier telles que celles décrites par Sambrook et al, Molecular cloning, Cold  
 15 Spring Harbor Laboratory Press, 1989. De telles conditions de stringence sont par exemple une hybridation à 65°C, pendant 18 heures dans une solution 5 x SSPE ; 10 x Denhardt's ; 100 µg/ml ADNss ; 1 % SDS suivie de 3 lavages pendant 5 minutes avec 2 x SSC ; 0,05 % SDS, puis 3 lavages pendant 15  
 20 minutes à 65°C dans 1 x SSC ; 0,1 % SDS. Les conditions de forte stringence comprennent par exemple une hybridation à 65°C, pendant 18 heures dans une solution 5 x SSPE ; 10 x Denhardt ; 100 µg/ml ADNss ; 1 % SDS suivie de 2 lavages pendant 20 minutes avec une solution 2 x SSC ; 0,05 % SDS à  
 25 65°C suivis d'un dernier lavage pendant 45 minutes dans une solution 0,1 x SSC ; 0,1 % SDS à 65°C. Les conditions de stringence moyenne comprennent par exemple un dernier lavage pendant 20 minutes dans une solution 0,2 x SSC, 0,1 % SDS à 65°C.

30 Par séquences qui présentent des homologies significatives, on inclut les séquences ayant une identité modérée ou importante de séquence nucléotidique avec l'une des séquences d'ADN ci-dessus et qui codent pour une protéine ayant la même fonction.

35 Par séquence d'ADN similaires, on entend ainsi des séquences d'ADN qui peuvent appartenir à d'autres mycètes que *Candida albicans* et notamment à *S.c.* et qui sont similaires ou identiques aux séquences d'ADN des gènes de *Candida*

*albicans* tels que définis ci-dessus. Ces séquences d'ADN similaires ne sont pas forcément identiques aux séquences d'ADN des gènes tels que définis ci-dessus. L'homologie de séquence au niveau nucléotidique peut-être modérée ou

5 importante. La présente invention concerne ainsi notamment les séquences d'ADN qui présentent une homologie de séquence nucléotidique d'au moins 50 %, de façon préférée d'au moins 60 % et de façon encore plus préférée d'au moins 70 % avec les séquences des gènes de la présente invention.

10 De plus, ces séquences d'ADN similaires ne codent pas forcément pour des protéines identiques, au niveau des séquences en acides aminés aux protéines codées par les gènes tels que définis ci-dessus. Ainsi la présente invention concerne notamment les séquences d'ADN qui codent pour des  
15 protéines dites homologues ayant une homologie de séquence en acides aminés d'au moins 40 %, notamment 45 %, de façon préférée au moins de 50 %, de façon plus préférée au moins de 60 % et de façon encore plus préférée au moins de 70 % avec les protéines codées par les gènes de la présente invention.

20 Chaque gène de la présente invention est représenté comme une séquence ADN simple brin comme indiqué dans SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 et SEQ ID N° 13 représentées respectivement dans le listing de séquences ci-après, mais il est entendu que la  
25 présente invention inclut la séquence ADN complémentaire de cette séquence ADN simple brin et inclut également la séquence ADN dite double brin constituée de ces deux séquences ADN complémentaires d'une de l'autre.

Les séquences d'ADN telles que définies ci-dessus sont  
30 des exemples de combinaison de codons codant pour les acides aminés correspondant respectivement aux séquences d'acides aminés SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14, telles que définies ci-dessus, mais il est entendu également que la  
35 présente invention inclut toute autre combinaison arbitraire de codons codant pour ces mêmes séquences d'acides aminés.

Pour la préparation des polynucléotides et notamment des séquences d'ADN telles que définies ci-dessus, des séquences



d'ADN modifiées comme indiqué ci-dessus ou encore des séquences d'ADN homologues telles que définies ci-dessus, on peut utiliser les techniques connues de l'homme du métier et notamment celles décrites dans l'ouvrage de Sambrook, J.

- 5 Fritsh, E. F. & Maniatis, T. (1989) intitulé : 'Molecular cloning : a laboratory manual', Laboratory, Cold Spring Harbor NY.

Les séquences d'ADN homologues telles que définies ci-dessus peuvent notamment être isolées selon les méthodes  
 10 connues de l'homme du métier par exemple par la technique de PCR en utilisant des amorces nucléotidiques dégénérées pour amplifier ces ADN à partir de banques génomiques ou de banques d'ADNc des mycètes correspondants. Les ADNc peuvent également être préparés à partir d'ARNm isolés de mycètes  
 15 d'espèces différentes étudiées dans le cadre de la présente invention telles que *Candida albicans* mais par exemple et tout aussi bien : *Candida stellatoidea*, *Candida tropicalis*, *Candida parapsilosis*, *Candida krusei*, *Candida pseudotropicalis*, *Candida quilliermondii*, *Candida glabrata*, *Candida*  
 20 *lusitanae* ou *Candida rugosa* ou encore des mycètes telles que *Saccharomyces cerevisiae* ou encore des mycètes du type *Aspergillus* ou *Cryptococcus* et notamment, par exemple, *Aspergillus fumigatus*, *Coccidioides immitis*, *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Blastomyces dermatitidis*, *Paracoc-*  
 25 *cidioides brasiliensis* et *Sporothrix schenckii* ou encore des mycètes des classes des phycomycètes ou eumycètes en particulier les sous-classes de basidiomycètes, ascomycètes, mebiascomycétales (levure) et plectascales, gymnascales (champignon de la peau et des cheveux) ou de la classe des  
 30 hyphomycètes, notamment les sous-classes conidiosporales et thallosporales parmi lesquels les espèces suivantes : *mucor*, *rhizopus*, *coccidioides*, *paracoccidioides* (*blastomyces*, *brasiliensis*), *endomyces* (*blastomyces*), *aspergillus*, *menicium* (*scopulariopsis*), *trichophyton* (*ctenomyces*), *epidermo-*  
 35 *phton*, *microsporon*, *piedraia*, *hormodendron*, *phialophora*, *sporotrichon*, *cryptococcus*, *candida*, *geotrichum*, *trichosporon* ou encore *toropsulosis*.

Les polynucléotides de la présente invention peuvent

ainsi être obtenus en utilisant les méthodes usuelles de clonage et de criblage telles que celles de clonage et séquençage à partir de fragments d'ADN chromosomique extraits de cellules ou encore issus de banques de gènes. Par exemple, 5 pour obtenir les polynucléotides de la présente invention, on peut partir d'une banque de fragments d'ADN chromosomique. On peut préparer une sonde correspondant à un oligonucléotide marqué par un élément radioactif, constituée de préférence de 17 nucléotides ou encore 20 ou plus et dérivée d'une séquence 10 partielle. Les clones contenant un ADN identique à celui de la sonde peuvent être ainsi identifiés sous des conditions stringentes. Par le séquençage de clones individuels ainsi identifiés, en utilisant des amorces de séquençage issues de la séquence d'origine, il est alors possible de prolonger la 15 séquence dans les deux directions pour déterminer la séquence du gène complet. De façon usuelle et efficace, un tel séquençage peut être réalisé en utilisant un ADN double brin dénaturé préparé à partir d'un plasmide. De telles techniques sont décrites par Maniatis, T. Fritsch, E.F. et Sambrook 20 comme indiqué ci-dessus. (Laboratory Manual, Cold Spring Harbor, New York (1989) (notamment en 1.90 et 13.70 dans les chapitres de screening par hybridation et séquençage à partir d'ADN double brin dénaturé).

Dans le cadre de la présente invention, on pourrait 25 notamment utiliser une banque de fragments d'ADN chromosomique de *Candida albicans* comme indiqué ci-après dans les exemples décrits dans la partie expérimentale.

Une description détaillée des conditions opératoires dans lesquelles a été réalisée la présente invention est 30 donnée ci-après.

L'invention a tout particulièrement pour objet les polypeptides ayant chacun une séquence d'acides aminés choisie parmi SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14, codés par 35 les séquences d'ADN telles que définies ci-dessus et les analogues de ces polypeptides.

Par analogues de polypeptides, on entend les polypeptides dont la séquence d'acides aminés a été modifiée

par substitution, suppression ou addition d'un ou plusieurs acides aminés mais qui conservent la même fonction biologique. De tels polypeptides analogues peuvent être produits spontanément ou peuvent être produits par

5 modification post-transcriptionnelle ou encore par modification de la séquence ADN de la présente invention comme indiqué ci-dessus, en utilisant les techniques connues de l'homme du métier : parmi ces techniques, on peut citer notamment la technique de mutagénèse dirigée connue de

10 l'homme du métier (Kramer, W., et al., Nucl. Acids Res., 12, 9441 (1984) ; Kramer, W. and Fritz, H.J., Methods in Enzymology, 154, 350 (1987) ; Zoller, M.J. and Smith, M. Methods in Enzymology, 100, 468 (1983)).

La synthèse d'ADN modifiés peut être faite comme

15 indiqué ci-dessus et notamment en utilisant des techniques de synthèse chimique bien connues telles que par exemple la méthode au phosphotriester [Letsinger, R.L and Ogilvie, K.K., K. Am. CHEM. Soc., 91, 3350 (1969) ; Merrifield, R.B., Sciences, 150, 178 (1968)] ou la méthode à la phosphoamidite

20 [Beaucage, S.L and Caruthers, M .H., Tetrahedron Lett., 22, 1859 (1981) ; McBRIDE, L.J. and Caruthers, M.H. Tetrahedron Lett., 24 245 (1983)] ou encore par la combinaison de ces méthodes.

Les polypeptides de la présente invention peuvent donc

25 être préparés par les techniques connues de l'homme du métier, notamment partiellement par synthèse chimique ou encore par la technique de l'ADN recombinant par expression dans une cellule hôte procaryote ou eucaryote comme indiqué ci-après.

30 La présente invention a particulièrement pour objet le procédé de préparation de protéines recombinantes PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 ayant respectivement les séquences d'acides aminés SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10,

35 SEQ ID N° 12 et SEQ ID N° 14, telles que définies ci-dessus, comprenant, pour la préparation de chacune de ces protéines, l'expression dans un hôte approprié de la séquence d'ADN telle que définie ci-dessus codant pour cette protéine puis

l'isolement et la purification de ladite protéine recombinante.

Pour produire les polypeptides de la présente invention, on peut notamment utiliser les techniques de l'ADN

- 5 recombinant en utilisant les méthodes de génie génétique et de culture cellulaire connues de l'homme du métier. On peut ainsi procéder par les étapes suivantes : d'abord préparation du gène approprié, puis incorporation de ce gène dans un vecteur, transfert du vecteur porteur du gène dans une  
10 cellule hôte appropriée, production du polypeptide par expression du gène, isolement du polypeptide, le polypeptide ainsi produit pouvant être ensuite purifié.

- Les polypeptides de la présente invention obtenus par l'expression des polynucléotides de la présente invention  
15 peuvent être purifiés à partir de cultures de cellules transformées par les méthodes bien connues de l'homme du métier telles que précipitation au sulfate d'ammonium ou à l'éthanol, extraction en conditions acides, chromatographie échangeuse d'anions ou de cations, chromatographie  
20 d'interaction hydrophobique, chromatographie d'affinité, chromatographie à l'hydroxylapatite et la chromatographie à haute performance liquide (HPLC). Des techniques bien connues de l'homme du métier peuvent être utilisées pour régénérer la protéine lorsque celle-ci est dénaturée durant son isolement  
25 ou sa purification.

- Les séquences d'ADN selon la présente invention et notamment SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 et SEQ ID N° 13 peuvent être préparées selon les techniques connues de l'homme du métier  
30 notamment par synthèse chimique ou par criblage d'une banque génomique ou d'une banque d'ADNc à l'aide de sondes d'oligonucléotides de synthèse par les techniques connues d'hybridation, ainsi amplification d'ADN à partir de fragments isolés ou encore par reverse transcriptase à partir  
35 d'ARN messager (ARNm).

L'avantage de la technique comprenant d'abord l'isolement d'ARNm par extraction des ARN totaux puis la synthèse d'ADNc à partir de ces ARNm par reverse

transcriptase réside notamment dans le fait que l'ARNm ne contient pas les introns alors que ces séquences non codantes peuvent être présentes dans l'ADN génomique.

On peut procéder en utilisant les techniques usuelles de  
5 clonage connues de l'homme du métier et notamment décrites dans l'ouvrage de Sambrook, J. Fritsh, E. F. & Maniatis, T. (1989) intitulé : 'Molecular cloning : a laboratory manual', Laboratory, Cold Spring Harbor NY.

Dans ces techniques, on peut procéder au clonage par  
10 insertion de fragment dans un plasmide qui peut être fourni avec un kit commercial adapté puis transformation d'une souche bactérienne par le plasmide ainsi obtenu. On peut utiliser notamment la souche *E. coli* XL1 Blue ou DH5 alpha. Les clones peuvent ensuite être cultivés pour extraire l'ADN  
15 plasmidique selon les techniques classiques de l'homme du métier référencées ci-dessus (Sambrook, Fritsh et Maniatis). On peut procéder au séquençage de l'ADN du fragment amplifié contenu dans l'ADN plasmidique.

Les polypeptides de la présente invention peuvent être  
20 obtenus par expression dans une cellule hôte contenant un polynucléotide selon la présente invention et notamment une séquence d'ADN codant pour un polypeptide de la présente invention précédée d'une séquence promoteur convenable. La cellule hôte peut être une cellule procaryote, par exemple *E.*  
25 *coli* ou une cellule eucaryote telle que les levures comme par exemple les Ascomycètes parmi lesquels les *Saccharomyces* ou encore des cellules de mammifères comme par exemple des cellules Cos.

La présente invention a particulièrement pour objet les  
30 vecteurs d'expression contenant pour chacun l'une des séquences d'ADN de la présente invention telles que définies ci-dessus.

Dans chacun de ces vecteurs d'expression, une telle  
séquence d'ADN est donc ainsi notamment la séquence d'ADN  
35 d'un gène de la présente invention codant pour une protéine de *Candida albicans* et contenant une séquence de nucléotides choisie parmi SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 et SEQ ID N° 13.

Dans chacun de ces vecteurs d'expression, une telle séquence d'ADN est ainsi encore plus particulièrement celle des gènes tels que définis ci-dessus codant pour l'une des séquences d'acides aminés SEQ ID N° 2, SEQ ID N° 4, SEQ ID  
 5 N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14 telles que définies ci-dessus et ci-après.

Dans chacun des vecteurs d'expression de la présente invention, une telle séquence d'ADN est ainsi une séquence d'ADN telle que définie ci-dessus codant pour l'une des  
 10 protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 ainsi que les séquences d'ADN qui hybrident avec celle-ci et/ou présentent des homologies significatives avec cette séquence ou des fragments de celle-ci ou encore les séquences d'ADN comprenant des modifications  
 15 introduites par suppression, insertion et/ou substitution d'au moins un nucléotide codant pour une protéine ayant la même activité.

Dans chacun des vecteurs d'expression de la présente invention, une telle séquence d'ADN est notamment une  
 20 séquence d'ADN telle que définie ci-dessus ainsi que les séquences d'ADN similaires qui ont une homologie de séquence nucléotidique d'au moins 50 % ou au moins 60 % et de préférence au moins 70 % avec ladite séquence d'ADN ou encore les séquences d'ADN similaires qui codent pour une protéine  
 25 dont la séquence en AA a une homologie d'au moins 40 % et notamment de 45 % ou d'au moins 50 %, plutôt au moins 60 % et de préférence au moins 70 % avec la séquence en AA codée par ladite séquence d'ADN.

Les vecteurs d'expression sont des vecteurs permettant  
 30 l'expression de la protéine sous le contrôle d'un promoteur convenable. Un tel vecteur peut être un plasmide, un cosmide ou un ADN viral. Pour les cellules procaryotes, le promoteur peut être par exemple le promoteur lac, le promoteur trp, le promoteur tac, le promoteur  $\beta$ -lactamase ou le promoteur PL.  
 35 Pour les cellules de levure, le promoteur peut être par exemple le promoteur PGK ou le promoteur GAL. Pour les cellules de mammifères, le promoteur peut être par exemple le promoteur SV40 ou les promoteurs de l'adénovirus.

Des vecteurs type Baculovirus peuvent être aussi utilisés pour l'expression dans des cellules d'insectes.

Les cellules hôtes sont par exemple des cellules procaryotes ou des cellules eucaryotes. Les cellules  
5 procaryotes sont par exemple *E. coli*, *Bacillus* ou *Streptomyces*. Les cellules hôtes eucaryotes comprennent des levures ainsi que des cellules d'organismes supérieurs, par exemple des cellules de mammifères ou des cellules  
10 d'insectes. Les cellules de mammifères sont par exemple des cellules CHO ou BHK de hamster ou des cellules Cos de singe. Les cellules d'insectes sont par exemple des cellules SF9.

La présente invention concerne donc un procédé qui comprend l'expression d'un polynucléotide selon la présente invention codant pour l'une des protéines PCaDR472, PCaDR489,  
15 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 dans une cellule hôte transformée par un polynucléotide selon la présente invention et notamment une séquence d'ADN codant pour la séquence en acides aminés SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ  
20 ID N° 14. Dans la réalisation d'un tel procédé, la cellule hôte est notamment une cellule eucaryote.

Pour la réalisation de la présente invention, les vecteurs utilisés peuvent être par exemple pGEX ou pBAD et la cellule hôte peut être *E. coli* ou par exemple le vecteur  
25 pYX222 et la cellule hôte peut être notamment *Saccharomyces cerevisiae*.

La présente invention a notamment pour objet la cellule hôte transformée avec un vecteur tel que défini ci-dessus et renfermant une séquence d'ADN selon la présente invention.

30 La présente invention a ainsi pour objet le procédé de préparation d'une protéine recombinante selon la présente invention, tel que défini ci-dessus, dans lequel la cellule hôte est *E. coli* DH5 alpha ou *E. coli* XL1-Blue ou notamment *Saccharomyces cerevisiae*.

35 Un exposé détaillé des conditions dans lesquelles peuvent être menées les opérations indiquées ci-dessus est donné ci-après dans la partie expérimentale. On a ainsi obtenu un plasmide dans lequel est inséré le gène de la

présente invention et on obtient ainsi également ce plasmide introduit dans une cellule hôte en opérant selon les techniques usuelles connues de l'homme du métier.

La présente invention a très précisément pour objet les 5 7 plasmides déposés le 25 mai 1999 à la Collection Nationale de Cultures de Microorganismes (CNCM) - INSTITUT PASTEUR - 25, rue du Docteur Roux - 75724 PARIS Cedex 15 sous les numéros suivants: I-2214, I-2215, I-2216, I-2217, I-2211, I-2212 et I-2213.

10 I-2214 est le numéro d'enregistrement de la souche CaDR472.10 constituée par la bactérie *E. coli* XL1-blue contenant un plasmide portant le gène de *Candida albicans* CaDR472 de la présente invention préparé comme indiqué à l'exemple 1 de la présente invention.

15 Ce gène correspond donc à la séquence CaDR472 de SEQ ID N° 1.

I-2215 est le numéro d'enregistrement de la souche CaDR489.37 constituée par la bactérie *E. coli* XL1-blue contenant un plasmide portant le gène de *Candida albicans* 20 CaDR489 de la présente invention préparé comme indiqué à l'exemple 2 de la présente invention.

Ce gène correspond donc à la séquence CaDR489 de SEQ ID N° 3.

I-2216 est le numéro d'enregistrement de la souche 25 CaDR527.2 constituée par la bactérie *E. coli* XL1-blue contenant un plasmide portant le gène de *Candida albicans* CaDR527 (allèle 1) de la présente invention préparé comme indiqué à l'exemple 3 de la présente invention.

Ce gène correspond donc à la séquence 1CaDR527 de SEQ ID 30 N° 5.

I-2217 est le numéro d'enregistrement de la souche CaDR527.3 constituée par la bactérie *E. coli* XL1-blue contenant un plasmide portant le gène de *Candida albicans* CaDR527 (allèle 2) de la présente invention préparé comme 35 indiqué à l'exemple 3 de la présente invention.

Ce gène correspond donc à la séquence 2CaDR527 de SEQ ID N° 7.



I-2211 est le numéro d'enregistrement de la souche CaFL024.4 constituée par la bactérie *E. coli* XL1-blue contenant un plasmide portant le gène de *Candida albicans* CaFL024 de la présente invention préparé comme indiqué à l'exemple 4 de la présente invention.

Ce gène correspond donc à la séquence CaFL024 de SEQ ID N° 9.

I-2212 est le numéro d'enregistrement de la souche CaNL260.4 constituée par la bactérie *E. coli* XL1-blue contenant un plasmide portant le gène de *Candida albicans* CaNL260 de la présente invention préparé comme indiqué à l'exemple 5 de la présente invention.

Ce gène correspond donc à la séquence CaNL260 de SEQ ID N° 11.

I-2213 est le numéro d'enregistrement de la souche CaDR361.3 constituée par la bactérie *E. coli* XL1-blue contenant un plasmide portant le gène de *Candida albicans* CaDR361 de la présente invention préparé comme indiqué à l'exemple 6 de la présente invention.

Ce gène correspond donc à la séquence CaDR361 de SEQ ID N° 13.

La présente invention a ainsi très précisément pour objet l'un ou plusieurs des plasmides déposés sous les numéros I-2214, I-2215, I-2216, I-2217, I-2211, I-2212 et I-2213.

Les conditions opératoires dans lesquelles a été réalisée la présente invention sont décrites ci-après dans la partie expérimentale.

La présente invention a ainsi pour objet un procédé de criblage de produits antifongiques caractérisé en ce qu'il comprend une étape où l'on mesure l'activité de l'une des protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 telles que définies ci-dessus en présence de chacun des produits dont on souhaite déterminer les propriétés antifongiques et l'on sélectionne les produits ayant un effet inhibiteur sur cette activité.

En particulier, les gènes codant pour les protéines

PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 de la présente invention étant essentiels à la survie des cellules de *Candida albicans*, des substances inhibitrices de telles protéines PCaDR472, PCaDR489, 5 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 pourraient être utilisables comme agents antifongiques, soit en tant que médicaments soit sur le plan industriel.

Par exemple, pour cribler des substances antifongiques telles que des substances actives sur *Candida albicans*, on 10 mesure l'activité d'une protéine codée par un gène de la présente invention ou de l'un de ses homologues fonctionnels et met la protéine en présence de chacun des produits dont on souhaite déterminer les propriétés antifongiques et l'on sélectionne ainsi les produits ayant un effet inhibiteur sur 15 cette activité.

On peut effectuer un tel criblage en mesurant l'activité de l'une des protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 de la présente invention en présence d'activateurs ou d'inhibiteurs 20 potentiels à tester, par exemple par mesure *in vitro* dans un milieu réactionnel approprié.

L'activité des protéines de la présente invention peut également être mesurée *in vivo* par un test cellulaire approprié. Par exemple, l'activité de PCaDR472, PCaDR489, 25 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 peut être avantageusement mesurée dans des cellules d'un mutant de *Saccharomyces cerevisiae* transformées par l'un des gènes de la présente invention et n'exprimant pas la protéine homologue PYDR 472w, PYDR 489w, PYDR 577w, PYFL 024c, PYNL 30 260c et PYDR 361c de *Saccharomyces cerevisiae*.

L'invention englobe également l'utilisation d'un produit sélectionné comme indiqué ci-dessus pour ses propriétés inhibitrices d'une des protéines de la présente invention pour l'obtention d'un agent antifongique.

35 La présente invention sera mieux comprise à l'aide de la partie expérimentale qui suit et qui décrit le clonage des gènes CaDR472, CaDR489, 1CaDR527, 2CaDR527, CaFL024, CaNL260 et CaDR361 de la présente invention.

La présente invention a ainsi pour objet l'utilisation d'un produit sélectionné par le procédé de criblage de produits antifongiques tel que défini ci-dessus pour l'obtention d'un agent antifongique.

- 5 La présente invention a également pour objet l'utilisation des gènes de *Candida albicans* de la présente invention ou des protéines codées par ces gènes tels que définis ci-dessus pour la sélection de produits ayant des propriétés antifongiques tels que définis ci-dessus et
- 10 utilisés comme inhibiteurs des protéines de *Candida albicans* codées par ces gènes.

- La présente invention a également pour objet les compositions pharmaceutiques renfermant à titre de principe actif au moins un inhibiteur des protéines de *Candida*
- 15 *albicans* de la présente invention telles que définies ci-dessus.

De telles compositions peuvent notamment être utiles pour traiter les infections fongiques topiques et systémiques.

- 20 Les compositions pharmaceutiques indiquées ci-dessus peuvent être administrées par voie buccale, rectale, par voie parentérale ou par voie locale en application topique sur la peau et les muqueuses ou par injection par voie intraveineuse ou intramusculaire. Ces compositions peuvent être solides ou
- 25 liquides et se présenter sous toutes les formes pharmaceutiques couramment utilisées en médecine humaine comme, par exemple, les comprimés simples ou dragéifiés, les gélules, les granulés, les suppositoires, les préparations injectables, les pommades, les crèmes, les gels et les
- 30 préparations en aérosols ; elles sont préparées selon les méthodes usuelles. Le principe actif peut y être incorporé à des excipients habituellement employés dans ces compositions pharmaceutiques, tels que le talc, la gomme arabique, le lactose, l'amidon, le stéarate de magnésium, le beurre de
- 35 cacao, les véhicules aqueux ou non, les corps gras d'origine animale ou végétale, les dérivés paraffiniques, les glycols, les divers agents mouillants, dispersants ou émulsifiants, les conservateurs.

La posologie sera variable selon le produit utilisé, le sujet traité et l'affection en cause.

La présente invention a ainsi notamment pour objet l'utilisation des compositions telles que définies ci-dessus  
5 comme agents antifongiques.

La présente invention a encore pour objet une méthode d'induction d'une réponse immunologique chez un mammifère comprenant l'inoculation à ce mammifère d'un polypeptide selon la présente invention tel que défini ci-dessus ou un  
10 fragment de ce polypeptide ayant la même fonction de façon à produire un anticorps permettant de protéger l'animal contre la maladie.

La présente invention a ainsi pour objet des anticorps dirigés contre les polypeptides de la présente invention tels  
15 que définis ci-dessus ou contre un fragment de ces polypeptides ayant la même fonction et codés par les polynucléotides de la présente invention et notamment par une séquence d'ADN telle que définie ci-dessus.

Les polypeptides de la présente invention peuvent ainsi  
20 être utilisés comme immunogènes pour produire des anticorps immunospcifiques de ces polypeptides. Le terme anticorps utilisé désigne les anticorps aussi bien monoclonaux que polyclonaux, chimériques, simple chaîne, les anticorps non humains et les anticorps humains, aussi bien que les  
25 fragments Fab, incluant ainsi les produits d'une banque d'immunoglobulines Fab. Les anticorps générés contre les polypeptides de la présente invention peuvent être obtenus par administration des polypeptides de la présente invention ou de fragments portant des épitopes, leurs analogues ou  
30 encore des cellules à un animal, de préférence non humain, en utilisant des protocoles de routine pour la préparation d'anticorps monoclonaux. De tels anticorps peuvent être préparés par les méthodes bien connues dans ce domaine telles que celles décrites dans l'ouvrage Antibodies, Laboratory  
35 manuel Ed. Harbow et David Larre, Cold Spring Harbor laboratory Eds, 1988.

La présente invention a ainsi tout particulièrement pour objet un anticorps dirigé contre l'une quelconque des

protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 de la présente invention ou un fragment de cette protéine. Un tel fragment a notamment la même fonction que la protéine dont il est issu.

5 La présente invention a encore pour objet l'utilisation des gènes CaDR472, CaDR489, 1CaDR527, 2CaDR527, CaFL024, CaNL260 et CaDR361 de la présente invention ou des protéines codées par ces gènes tels que définis ci-dessus pour la  
10 préparation de compositions utiles pour le diagnostic ou le traitement de maladies causées par la levure pathogène *Candida albicans*.

La présente invention concerne aussi l'utilisation des polynucléotides de la présente invention comme réactifs de diagnostic. La détection d'un polynucléotide selon la  
15 présente invention codant pour l'une des protéines de *Candida albicans* de la présente invention ou de ses analogues chez un eucaryote en particulier un mammifère et plus particulièrement un être humain, peut constituer un moyen de diagnostic d'une maladie : ainsi, on peut détecter un tel  
20 polynucléotide selon la présente invention et notamment une séquence d'ADN par une grande variété de techniques chez un eucaryote en particulier un mammifère et plus particulièrement un être humain, infectés par un organisme contenant au moins l'un des polynucléotides de la présente  
25 invention. Les acides nucléiques pour une telle utilisation d'outil de diagnostic peuvent être détectés à partir de cellules ou de tissus infectés, tels que l'os, le sang, le muscle, le cartilage ou la peau. Pour cette détection, l'ADN génomique peut être utilisé directement ou encore être  
30 amplifié par PCR ou une autre technique d'amplification. Les ARN ou ADN et ADNc peuvent également être utilisés dans le même but. Par les techniques d'amplification, la lignée du mycète présent dans un eucaryote en particulier un mammifère et plus particulièrement un être humain, peut être  
35 caractérisée par l'analyse du génotype. Des délétions ou des insertions peuvent être détectées par le changement de taille du produit amplifié par comparaison avec le génotype de la séquence de référence. Les points de mutations peuvent être

identifiés par hybridation de l'ADN amplifié avec les séquences, marquées par un élément radioactif, de polynucléotides de la présente invention. Des séquences parfaitement complémentaires peuvent ainsi être distinguées de duplex qui résistent mal à la digestion par des nucléases. Les différences de séquences d'ADN peuvent aussi être détectées par des altérations de la mobilité électrophorétique de fragments d'ADN dans des gels, avec ou sans agent dénaturant, ou par un séquençage direct d'ADN (référence : Myers et al. Science, 230 : 1242 (1985)).

Des changements de séquences à des localisations spécifiques peuvent aussi être révélés par des expériences de protection contre des nucléases telles que RNase I et S1 ou par des méthodes de clivage chimique (référence : Cotton et al., Proc Natl Acad Sci, USA, 85 : 4397-4401 (1985)).

Des cellules contenant l'un des polynucléotides de la présente invention portant des mutations ou des polymorphismes peuvent aussi être détectées par un grand nombre de techniques permettant notamment de déterminer le sérotype. Par exemple, la technique RT-PCR peut être utilisée pour détecter les mutations. Il est particulièrement préféré d'utiliser les techniques de RT-PCR en conjonction avec des systèmes de détection automatique, tels que par exemple dans la technique GeneScan. ARN et ADNc peuvent être utilisés dans les techniques PCR ou RT-PCR. Par exemple, des amorces complémentaires des polynucléotides codant pour les polypeptides de la présente invention peuvent être utilisées pour identifier et analyser les mutations.

Des amorces peuvent ainsi être utilisées pour amplifier un ADN isolé de l'individu infecté. De cette façon des mutations dans la séquence d'ADN peuvent être détectées et utilisées pour diagnostiquer l'infection et déterminer le sérotype ou le classement de l'agent infectieux. De telles techniques sont usuelles pour l'homme du métier et sont décrites notamment dans le manuel 'Current Protocols in Molecular Biology, Ausubel et al, ed. John Wiley & sons, Inc., 1995).

La présente invention concerne ainsi un procédé de

diagnostic d'une maladie et de préférence d'une infection fongique provoquée par *Candida albicans* telles que des mycoses comme indiqué ci-dessus, ce procédé comprenant la détermination à partir d'un échantillon prélevé sur un  
5 individu infecté, d'une augmentation de la quantité de l'un des polynucléotides de la présente invention. Un tel polynucléotide peut notamment avoir l'une des séquences d'ADN de la présente invention telles que définies ci-dessus.

Des augmentations ou des diminutions de la quantité de  
10 polynucléotides peuvent être mesurées par les techniques bien connues de l'homme du métier telles que notamment l'amplification, la PCR, RT-PCR, Northern blotting ou autres techniques d'hybridation.

De plus, une méthode de diagnostic en accord avec la  
15 présente invention consiste en la détection d'une expression trop importante de polypeptides de la présente invention, par comparaison avec des échantillons de contrôle constitués de tissus normaux non infectés utilisés pour détecter la présence d'une infection.

20 Les techniques qui peuvent être utilisées pour détecter ainsi les quantités de protéines exprimées dans un échantillon d'une cellule hôte sont bien connues de l'homme du métier. On peut ainsi citer par exemple les techniques de radioimmunoassay ou de competitive-binding, analyse par  
25 Western Blot et test ELISA (ref Ausubel indiqué ci-dessus).

La présente invention a encore pour objet un kit pour le diagnostic d'infections fongiques comprenant une séquence d'ADN selon la présente invention telle que définie ci-dessus ou une séquence similaire ou un fragment de cette séquence,  
30 le polypeptide codé par cette séquence ou un fragment polypeptidique ayant la même fonction ou un anticorps dirigé contre un tel polypeptide codé par cette séquence d'ADN ou contre un fragment de ce polypeptide.

Ce kit pourra ainsi contenir une séquence d'ADN selon la  
35 présente invention telle que définie ci-dessus soit par exemple la séquence d'ADN SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 ou SEQ ID N° 13 ou un fragment de cette séquence.

Un tel kit pourra de même contenir un polypeptide selon la présente invention ou un fragment de ce polypeptide et notamment l'une des protéines selon la présente invention ayant la séquence en AA SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14 ou encore un anticorps tel que défini ci-dessus.

Un tel kit peut-être préparé selon les méthodes bien connues de l'homme du métier.

Le listing de séquences SEQ ID N° 1 à SEQ ID N° 32 et les figures 1 à 6 ci-après présentent les illustrations suivantes qui permettent de mieux décrire la présente invention.

Les séquences SEQ ID N° 1 à SEQ ID N° 32 représentent les séquences nucléotidiques ou peptidiques indiquées dans la présente invention.

Les séquences SEQ ID N° 1 à SEQ ID N° 14 décrivent les séquences nucléotidiques des gènes de *Candida albicans* de la présente invention et les séquences peptidiques des protéines déduites de ces gènes.

Les séquences SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 ou SEQ ID N° 13 décrivent ainsi respectivement les séquences nucléotidiques des gènes de *Candida albicans* de la présente invention : CaDR472, CaDR489, 1CaDR527, 2CaDR527, CaFL024, CaNL260 et CaDR361.

Les séquences SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14 décrivent respectivement les séquences peptidiques des protéines déduites des gènes de la présente invention.

Ainsi, par exemple, les séquences SEQ ID N° 1 et SEQ ID N° 2 représentent respectivement la séquence nucléotidique du gène CaDR472 et la séquence peptidique de la protéine déduite de ce gène soit PCaDR472.

Les séquences SEQ ID N° 15 à SEQ ID N° 20 représentent respectivement les séquences des 6 sondes utilisées pour la préparation des gènes de *Candida albicans* de la présente invention comme indiqué ci-après dans la partie expérimentale.

Les séquences SEQ ID N° 21 à SEQ ID N° 32 représentent



respectivement les séquences des 2 x 6 oligonucléotides utilisés pour amplifier les sondes pour la préparation des gènes de *Candida albicans* de la présente invention comme indiqué ci-après dans la partie expérimentale.

- 5 Les figures 1 à 6 ci-après se réfèrent chacune respectivement à une des 6 préparations des gènes de *Candida albicans* de la présente invention soit : CaDR472, CaDR489, 1CaDR527/2CaDR527, CaFL024, CaNL260 et CaDR361, ces préparations étant décrites ci-après à la partie  
10 expérimentale aux exemples 1 à 6.

Chacune des figures 1 à 6 décrit la comparaison de la protéine déduite de la sonde utilisée pour la préparation d'un des gènes de *Candida albicans* de la présente invention (les 6 sondes utilisées ayant les séquences SEQ ID N° 15 à  
15 SEQ ID N° 20) à la séquence du gène de *S.c.* pris comme point de départ de la préparation de ce gène de *Candida albicans*.

Ainsi, se référant à l'exemple 1 de préparation du gène CaDR472 de la présente invention, la figure 33 représente la comparaison de la protéine déduite de la sonde de CaDR472  
20 (SEQ ID N° 15) à la protéine déduite du gène YDR472w de *S. cerevisiae*.

La partie expérimentale ci-après permet de décrire la présente invention sans toutefois la limiter.

#### Partie expérimentale

- 25 **EXEMPLE 1 : Clonage et séquençage du gène CaDR472**  
(méthode A)

Le site Internet de Stanford  
(<http://candida.stanford.edu/>) permet d'accéder directement aux séquences préliminaires du génome de *Candida albicans*.

- 30 L'une de ces séquences présente une homologie avec le gène YDR472w de *S. cerevisiae*. Deux oligonucléotides ont été choisis dans cette séquence soit :

5'CAATTTATTC ATGTTTCGNAT CTGGAAATTG ATTTT3' nommé SEQ ID N° 21  
et 5'CCAAATCTCA AACTCTCTCT AATTAAAC3' nommé SEQ ID N° 22.

- 35 Ces deux oligonucléotides sont utilisés pour amplifier le fragment de *C. albicans*. Après clonage, une séquence dite sonde de CaDR472 de 320 paires de bases proche de la séquence attendue a été obtenue : la sonde de CaDR472 est appelée SEQ

ID NO 15. La protéine déduite de la sonde de CaDR472 (SEQ ID NO 15) a été comparée à celle de YDR472w ce qui met en évidence une identité de 48% entre ces deux séquences d'AA : cette comparaison est représentée à la figure 1.

- 5 Le fragment de 320 paires de bases de *C. albicans* a été utilisé comme sonde pour criblage de la banque génomique de *C. albicans* : cette banque de C.a. a été préparée par digestion partielle de l'ADN génomique de *C. albicans* par Sau3AI et clonage dans le vecteur YEP24 au site de
- 10 restriction BamHI. Les clones de la banque génomique ont ensuite été étalés à la densité de 2000 clones par boîte : chaque boîte est ensuite recouverte d'un filtre de nitrocellulose qui est successivement traité par : NaOH, 0.5M, pendant 5 minutes ; Tris, 1M, pH 7.7, pendant 5
- 15 minutes ; Tris, 0.5M, pH 7.7, NaCl, 1.25M, pendant 5 minutes. Après séchage, les filtres sont gardés pendant deux heures à 80°C. Préhybridation et hybridation sont réalisées dans un tampon de 40 % de formamide, 5xSSC, 20 mM Tris pH 7,7 1xDenhardt 0,1 % SDS. La sonde est ensuite marquée au P32 par
- 20 le kit Rediprime et dCTP 32p de Amersham UK. L'hybridation est réalisée en 17 heures à 42°C. Les filtres sont ensuite lavés au 1xSSC, 0,1 % SDS, trois fois pendant 5 minutes à la température ambiante et ensuite deux fois pendant 30 minutes à 60°C puis sont soumis à une autoradiographie
- 25 pendant une nuit. Les colonies correspondant aux spots obtenus sont isolées par un nouvel étalement à faible densité suivi d'hybridation : 8 clones positifs sont ainsi obtenus (à partir de 60 000) qui sont alors séquencés à l'aide d'un appareil ABI 377. Les séquences sont compilées à l'aide d'un
- 30 software ABI puis analysées à l'aide d'un package software GCG. L'un des 8 clones s'est révélé contenir la séquence codante complète correspondant à la sonde utilisée : ce gène est appelé CaDR472 et cette séquence est dite SEQ ID NO 1.

CaDR472 a 47.5 % de nucléotides identiques à YDR472w de  
35 *S. cerevisiae*.

Pour la traduction en acides aminés, il a été tenu compte du fait que dans *C. albicans* le codon CTG est traduit en sérine (il y a 1 codon CTG dans CaDR472). La protéine

déduite du gène CaDR472 (SEQ ID N° 1) soit SEQ ID N° 2 (PCaDR472) a 52,4 % de similarité en acides aminés et 44 % d'identité en acides aminés avec la protéine déduite de YDR472w.

- 5 La séquence complète du gène CaDR472 contient un codon CTG.

**EXEMPLE 2 : Clonage et séquençage du gène CaDR489**

On procède comme à l'exemple 1 à partir de séquences préliminaires du génome de *Candida albicans* du site Internet  
10 de Stanford (<http://candida.stanford.edu/>). L'une de ces séquences présente une homologie avec le gène YDR489w de *S. cerevisiae*. Deux oligonucléotides ont été choisis dans cette séquence soit :

- 5' GTTCATGTTT GGTGACTCAG AGCGTCTCAA CTATATTG3' nommé SEQ ID  
15 N° 23  
et 5' TTTGATAAAC ACAGGCTGGT CTAAATCTGG CTC3' nommé SEQ ID  
N° 24.

Ces deux oligonucléotides sont utilisés pour amplifier le fragment de *C. albicans*. Après clonage, une séquence dite  
20 sonde de CaDR489 de 295 paires de bases proche de la séquence attendue a été obtenue : la sonde de CaDR489 est appelée SEQ ID N° 16. La protéine déduite de la sonde de CaDR489 (SEQ ID N° 16) a été comparée à celle de YDR489w ce qui met en évidence une identité de 41% entre ces deux séquences d'AA :  
25 cette comparaison est représentée à la figure 2.

Le fragment de 295 paires de bases de *C. albicans* a été utilisé comme sonde pour criblage de la banque de gènes de *C. albicans* préparée par digestion partielle du DNA génomique de *C. albicans* en procédant comme à l'exemple 1.

- 30 Le clonage est réalisé comme indiqué à l'exemple 1 et après préhybridation et hybridation réalisées comme indiqué à l'exemple 1, on obtient 4 clones positifs (à partir de 40 000). On procède au séquençage et l'analyse des séquences obtenues comme indiqué à l'exemple 1, et on obtient ainsi un  
35 clone se révélant contenir la séquence codante complète correspondant à la sonde utilisée : ce gène est appelé CaDR489 et cette séquence est dite SEQ ID N° 4.  
CaDR489 a 48.1 % de nucléotides identiques à YDR489w de *S.*

*cerevisiae*.

La protéine déduite du gène CaDR489 (SEQ ID N° 3) soit SEQ ID n° 4 ou PcaDR489 a 50 % de similarité en acides aminés et 37 % d'identité en acides aminés avec la protéine déduite de YDR489.

La séquence complète du gène CaDR489 contient un codon CTG.

**EXEMPLE 3 : Clonage et séquençage du gène CaDR527**

On procède comme à l'exemple 1 à partir de séquences préliminaires du génome de *Candida albicans* du site Internet de Stanford (<http://candida.stanford.edu/>). L'une de ces séquences présente une homologie avec le gène YDR527w de *S. cerevisiae*. Deux oligonucléotides ont été choisis dans cette séquence soit :

5'ATCTCTGATA TGAGATTTGG CTTTAAAGGC GA3' nommé SEQ ID N° 25 et 5'GGTCTTTTTT CCATCAGCTG CCTCTGTTAT TG3' nommé SEQ ID N° 26.

Ces deux oligonucléotides sont utilisés pour amplifier le fragment de *C. albicans*. Après clonage, une séquence dite sonde de CaDR527 de 392 paires de bases proche de la séquence attendue a été obtenue : la sonde de CaDR527 est appelée SEQ ID N° 17. La protéine déduite de la sonde de CaDR527 (SEQ ID N° 17) a été comparée à celle de YDR527w ce qui met en évidence une identité de 41% entre ces deux séquences d'AA : cette comparaison est représentée à la figure 3.

Le fragment de 392 paires de bases de *C. albicans* a été utilisé comme sonde pour le criblage de la banque génomique de *C. albicans* préparée par digestion partielle du DNA génomique de *C. albicans* en procédant comme à l'exemple 1.

Le clonage est réalisé comme indiqué à l'exemple 1 et après préhybridation et hybridation réalisées comme indiqué à l'exemple 1, on obtient 7 clones positifs (à partir de 40 000). On procède au séquençage et l'analyse des séquences obtenues comme indiqué à l'exemple 1.

On obtient ainsi deux clones se révélant contenir chacun une séquence codante complète correspondant chacune à un allèle de la sonde utilisée : ce gène est appelé CaDR527 et les deux allèles sont ainsi appelés 1CaDR527 et 2CaDR527 et

leurs séquences respectives sont respectivement dites SEQ ID N° 5 et SEQ ID N° 7.

On constate que les gènes des allèles 1CaDR527 et 2CaDR527 (SEQ ID N° 5 et SEQ ID N° 7) diffèrent par 13 5 nucléotides.

Le gène CaDR527 (1er allèle) a 53.8 % de nucléotides identiques à YDR527w de *S. cerevisiae*.

Les protéines déduites de ces allèles soit SEQ ID N° 6 (PCaDR527) pour le 1er allèle 1CaDR527 et SEQ ID N° 8 pour le 10 2ème allèle 2CaDR527 diffèrent entre elles par 5 acides aminés.

La protéine déduite du gène CaDR527 (SEQ ID N° 6) a 58,9 % de similarité en acides aminés et 47,9 % d'identité en acides aminés avec la protéine déduite de YDR527.

15 La séquence complète du gène CaDR527 ne contient pas de codon CTG.

**EXEMPLE 4 : Clonage et séquençage du gène CaFL024**  
(méthode B)

Le site Internet de Stanford (<http://candida.stanford.edu/>) 20 permet d'accéder directement aux séquences préliminaires du génome de *Candida albicans*. L'une de ces séquences présente une homologie avec le gène YFL024c de *S. cerevisiae*. Deux oligonucléotides ont été choisis dans cette séquence soit :

5' ATTCCCACAC CGGACGCTTC 3' nommé SEQ ID N° 27  
25 et 5'GACAACTCCT CGTACGATAG 3' nommé SEQ ID N° 28.

Ces deux oligonucléotides sont utilisés pour amplifier le fragment de *C. albicans*. Après clonage, une séquence dite sonde de CaFL024 de 335 paires de bases proche de la séquence attendue a été obtenue : la sonde de CaFL024 est appelée SEQ 30 ID N° 18. La protéine déduite de la sonde de CaFL024 (SEQ ID N° 18) a été comparée à celle de YFL024c ce qui met en évidence une similarité de 62 % et une identité de 58 % entre ces deux séquences d'AA : cette comparaison est représentée à la figure 4.

35 Ce fragment de 335 paires de bases de *C. albicans* a été utilisé comme sonde pour criblage d'une banque génomique de *C. albicans* : cette banque de gènes de *C.a.* a été préparée par digestion partielle du DNA génomique de *C. albicans* par

SauIIIA et clonage dans le vecteur YEP-24 au site de restriction BamHI. Les clones de la banque de gènes ont ensuite été étalés à la densité de 2000 clones par boîte : chaque boîte est ensuite recouverte d'un filtre de

5 nitrocellulose qui est successivement traité par : 1.5 M NaCl/ 0.5 M NaOH pendant 5 minutes; 1.5 M NaCl/0.5 M Tris-HCl pH 7.2/1 mM EDTA pendant 3 minutes, à deux reprises.

Le DNA est ensuite 'crosslinked' aux filtres (Amersham Life Science, ultraviolet crosslinker).

10 La sonde (100 ng) est ensuite marquée au P32 par le kit Rediprime et dCTP (Amersham Life Science).

Préhybridation et hybridation des filtres sont réalisées dans un tampon de 30 % de formamide, 5 x SSC, 5 % de solution de Denhardt, 1 % SDS, 100 µg/ml de DNA de sperme de saumon et

15 une concentration de la sonde de 10(6) cpm/ml : l'hybridation est réalisée à 42°C pendant 16 heures.

Les filtres sont ensuite lavés trois fois, pendant 5 minutes chaque fois, à la température ambiante au 2 x SSC/ 0,1 % SDS puis trois fois au 1 x SSC/ 0,1 % SDS pendant 20

20 minutes chaque fois à 60°C. Les filtres sont soumis à une autoradiographie pendant une nuit. Les colonies correspondant aux clones positifs (spots obtenus) sont isolées et soumis à un second criblage par un nouvel étalement à faible densité suivi d'hybridation : 6 clones sont ainsi obtenus (à partir

25 de 144 000) qui sont alors séquencés à l'aide d'un appareil ABI 377. Les séquences sont compilées à l'aide d'un software ABI puis analysées à l'aide d'un package software GCG. L'un des 6 clones s'est révélé contenir la séquence codante complète correspondant à la sonde utilisée : ce gène est

30 appelé CaFL024 et cette séquence dite SEQ ID NO 9.

CaFL024 a 49.1 % de nucléotides identiques à YFL024c de *S. cerevisiae*.

Il y a 2 codons CTG dans CaFL024. La protéine déduite du gène CaFL024 (SEQ ID N° 9) soit SEQ ID n° 10 (PCaFL024) a

35 51,8 % de similarité en acides aminés et 44,0 % d'identité en acides aminés avec la protéine déduite de YFL024c.

#### **EXEMPLE 5 : Clonage et séquençage du gène CaNL260**

On procède comme à l'exemple 4 à partir de séquences

préliminaires du génome de *Candida albicans* du site Internet de Stanford (<http://candida.stanford.edu/>). L'une de ces séquences présente une homologie avec le gène YNL260c de *S. cerevisiae*. Deux oligonucléotides ont été choisis dans cette

5 séquence soit :

5' AGATAATGTATTAAATTTAG 3' nommé SEQ ID N° 29

et 5' CTCTTAATTTATTTCTTGCC 3' nommé SEQ ID N° 30.

Ces deux oligonucléotides sont utilisés pour amplifier le fragment de *C. albicans*. Après clonage, une séquence dite

10 sonde de CaNL260 de 326 paires de bases proche de la séquence attendue a été obtenue : la sonde de CaNL260 est appelée SEQ ID N° 19. La protéine déduite de la sonde de CaNL260 (SEQ ID N° 19) a été comparée à celle de YNL260c ce qui met en évidence une similarité de 56,7 % et une identité de 40,3 %

15 entre ces deux séquences d'AA : cette comparaison est représentée à la figure 5.

Le fragment de 326 paires de bases de *C. albicans* a été utilisé comme sonde pour criblage de la banque de gènes de *C. albicans* préparée par digestion partielle du DNA génomique de

20 *C. albicans* en procédant comme à l'exemple 4.

La préhybridation et hybridation sont réalisés comme indiqué à l'exemple 4, on obtient 2 clones positifs (à partir de 40 000). On procède au séquençage et l'analyse des séquences obtenues comme indiqué à l'exemple 4, et on obtient

25 ainsi un clone se révélant contenir la séquence codante complète correspondant à la sonde utilisée : ce gène est appelé CaNL260 et cette séquence est dite SEQ ID N° 11.

CaNL260 a 47.6 % de nucléotides identiques à YNL260c de *S. cerevisiae*.

30 La protéine déduite du gène CaNL260 (SEQ ID N° 11) soit SEQ ID N° 12 (PCaNL260) a 50,7 % de similarité en acides aminés et 32,6 % d'identité en acides aminés avec la protéine déduite de YNL260c.

Il n'y a pas de codon CTG dans CaNL260.

### 35 EXEMPLE 6 : Clonage et séquençage du gène CaDR361

On procède comme à l'exemple 4 à partir de séquences préliminaires du génome de *Candida albicans* :

Le site Internet de Stanford (<http://candida.stanford.edu/>)

permet d'accéder directement aux séquences préliminaires du génome de *Candida albicans*.

L'une de ces séquences présente une homologie avec le gène YDR361c de *S. cerevisiae*. Deux oligonucléotides ont été  
5 choisis dans cette séquence soit :

5' CCTCAAATTGATTTCCATGC 3' nommé SEQ ID N° 31

et 5'GTGGAATCACTTCAACTGGC 3' nommé SEQ ID N° 32.

Ces deux oligonucléotides sont utilisés pour amplifier le fragment de *C. albicans*. Après clonage, une séquence dite  
10 sonde de CaDR361 de 374 paires de bases proche de la séquence attendue a été obtenue : la sonde de CaDR361 est appelée SEQ ID N° 20. La protéine déduite de la sonde de CaDR361 (SEQ ID N° 20) a été comparée à celle de YDR361c ce qui met en évidence une similarité de 52,4 % et une identité de 40,0 %  
15 entre ces deux séquences d'AA : cette comparaison est représentée à la figure 6.

Le fragment de 374 paires de bases de *C. albicans* a été utilisé comme sonde pour criblage de la banque de gènes de *C. albicans* préparée par digestion partielle du DNA génomique de  
20 *C. albicans* par Sau11/A et clonage dans le vecteur YEP 24 (marqueur de sélection Trp) au site de restriction Bam HI.

La préhybridation et hybridation réalisés comme indiqué à l'exemple 4, on obtient 4 clones positifs (à partir de 40 000). On procède au séquençage et l'analyse des séquences  
25 obtenues comme indiqué à l'exemple 4, et on obtient ainsi un clone se révélant contenir la séquence codante complète correspondant à la sonde utilisée : ce gène est appelé CaDR361 et cette séquence dite SEQ ID N° 13.

CaDR361 a 53.9 % de nucléotides identiques à YDR361c de  
30 *S. cerevisiae*.

CaDR361I1 n'y a pas de codon CTG dans CaDR361.



REVENDICATIONS

- 1) Polynucléotides isolés contenant chacun une séquence nucléotidique choisie dans le groupe suivant:
  - 5 a) un polynucléotide ayant au moins 50 % ou au moins 60 % et de préférence au moins 70 % d'identité avec un polynucléotide codant pour un polypeptide ayant la même fonction et ayant une séquence en acides aminés homologue d'une séquence choisie parmi SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14
  - b) un polynucléotide complémentaire du polynucléotide a)
  - c) un polynucléotide comprenant au moins 15 bases consécutives du polynucléotide défini en a) et b).
- 2) Polynucléotides selon la revendication 1 tels que ces
  - 15 polynucléotides sont des ADN.
- 3) Polynucléotides selon la revendication 1 tels que ces polynucléotides sont des ARN.
- 4) Polynucléotides tels que définis à la revendication 2 comprenant chacun une séquence de nucléotides choisie parmi
  - 20 SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 et SEQ ID N° 13.
- 5) Séquences d'ADN telles que définies aux revendications 1, 2 et 4 caractérisées en ce que ces séquences d'ADN sont celles des gènes codant respectivement pour des protéines de
  - 25 Candida albicans (ayant les mêmes fonctions que les protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361) et contenant chacune une séquence de nucléotides choisie parmi SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 et SEQ ID N° 13.
- 6) Séquences d'ADN de gènes selon la revendication 5 ou 6 codant chacune pour une séquence d'acides aminés choisie parmi SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14.
- 7) Séquences d'ADN codant pour les protéines PCaDR472,
  - 35 PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 selon les revendications 5 et 6 ainsi que les séquences d'ADN qui hybrident avec celles-ci et/ou présentent des homologies significatives avec ces séquences ou des fragments de celles-

ci et codent pour des protéines ayant les mêmes fonctions.

- 8) Séquences d'ADN selon les revendications 5 à 7 comprenant des modifications introduites par suppression, insertion et/ou substitution d'au moins un nucléotide codant pour des
  - 5 protéines ayant les mêmes activités que les protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361.
  - 9) Séquences d'ADN selon l'une des revendications 5 à 8 ainsi que les séquences d'ADN qui ont une homologie de séquence
    - 10 nucléotidique d'au moins 50 % ou au moins 60 % et de préférence au moins 70 % avec lesdites séquences d'ADN.
    - 10) Séquences d'ADN selon l'une des revendications 5 à 9 ainsi que les séquences d'ADN qui codent pour des protéines de fonctions similaires dont les séquences respectives en AA
      - 15 ont une homologie d'au moins 40 % et notamment de 45 % ou d'au moins 50 %, plutôt au moins 60 % et de préférence au moins 70 % avec les séquences en AA codées par lesdites séquences d'ADN.
      - 11) Polypeptides ayant chacun une séquence d'acides aminés
        - 20 choisie parmi SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14 codées par les séquences d'ADN selon l'une des revendications 5 à 10 et les analogues de ces polypeptides.
        - 12) Procédé de préparation de protéines recombinantes
          - 25 PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 ayant respectivement les séquences d'acides aminés SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14 comprenant, pour la préparation de chacune de ces protéines, l'expression dans un
            - 30 hôte approprié de la séquence d'ADN codant pour cette protéine selon l'une des revendications 5 à 10 puis l'isolement et la purification de ladite protéine recombinante.
            - 13) Vecteurs d'expression contenant pour chacun l'une des
              - 35 séquences d'ADN selon l'une des revendications 5 à 10.
              - 14) Cellule hôte transformée avec un vecteur selon la revendication 13.
              - 15) Procédé tel que défini à la revendication 12 dans lequel

la cellule hôte est E. coli DH5 alpha ou E. coli XL1-Blue.

16) Procédé tel que défini à la revendication 13 dans laquelle la cellule hôte est *Saccharomyces cerevisiae*.

17) L'un ou plusieurs des plasmides déposés à la CNCM sous  
5 les numéros I-2214, I-2215, I-2216, I-2217, I-2211, I-2212 et I-2213.

18) Procédé de criblage de produits antifongiques caractérisé en ce qu'il comprend une étape où l'on mesure l'activité de l'une des protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527,  
10 PCaFL024, PCaNL260, PCaDR361, telles que définies à la revendication 11, en présence de chacun des produits dont on souhaite déterminer les propriétés antifongiques et l'on sélectionne les produits ayant un effet inhibiteur sur cette activité.

15 19) Utilisation d'un produit sélectionné par le procédé selon la revendication 19 pour l'obtention d'un agent antifongique.

20) Utilisation des gènes de *Candida albicans* ou des protéines codées par ces gènes selon l'une des revendications 5 à 11 pour la sélection de produits ayant des propriétés  
20 antifongiques selon la revendication 19 comme inhibiteurs des protéines de *Candida albicans* codées par ces gènes.

21) Compositions pharmaceutiques renfermant à titre de principe actif au moins un inhibiteur des protéines de *Candida albicans* tel que défini à la revendication 20.

25 22) Utilisation des compositions telles que définies à la revendication 21 comme agents antifongiques.

23) Méthode d'induction d'une réponse immunologique chez un mammifère comprenant l'inoculation à ce mammifère d'un polypeptide tel que défini à la revendication 11 ou un  
30 fragment de ce polypeptide ayant la même fonction de façon à produire un anticorps permettant de protéger l'animal contre la maladie.

24) Anticorps dirigé contre un polypeptide tel que défini à la revendication 11 ou un fragment de ce polypeptide ayant la  
35 même fonction.

25) Anticorps tel que défini à la revendication 24 dirigé contre l'une quelconque des protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 ou un

fragment de ces protéines.

- 26) Utilisation de l'un quelconque des gènes CaDR472, CaDR489, 1CaDR527, 2CaDR527, CaFL024, CaNL260 et CaDR361 ou de l'une quelconque des protéines codées par ces gènes selon
- 5 l'une des revendications 5 à 11 pour la préparation de compositions utiles pour le diagnostic ou le traitement de maladies causées par la levure pathogène *Candida albicans*.
- 27) Kit pour le diagnostic d'infections fongiques comprenant une séquence d'ADN tel que défini à l'une des revendications
- 10 5 à 11 ou une séquence ayant une fonction similaire ou un fragment fonctionnel de cette séquence, le polypeptide codé par cette séquence ou un fragment polypeptidique ayant la même fonction ou un anticorps dirigé contre un tel polypeptide codé par cette séquence d'ADN ou contre un
- 15 fragment de ce polypeptide.

## LISTE DE SEQUENCES

<110> Hoechst Marion Roussel

<120> Gènes de Candida albicans et les protéines codées par  
ces gènes.

<130> 9902

<140>

<141>

<160> 32

<170> PatentIn Ver. 2.1

<210> 1

<211> 747

<212> ADN

<213> Candida albicans

<220>

<221> CDS

<222> (1)..(747)

<220>

<221> modified\_base

<222> (136)..(138)

<400> 1

atg tca aat gac gat ata ata ctc cca tca gtt tca tcc tta tcg aaa 48

Met Ser Asn Asp Asp Ile Ile Leu Pro Ser Val Ser Ser Leu Ser Lys

1

5

10

15

cta act ata aat gat gta tca aaa tca gga ttt gga tac aat ccg tcc 96

Leu Thr Ile Asn Asp Val Ser Lys Ser Gly Phe Gly Tyr Asn Pro Ser

20

25

30

ata gga cca ata tca aat act att acc cta gaa tct tca ctg gta tta	144
Ile Gly Pro Ile Ser Asn Thr Ile Thr Leu Glu Ser Ser Ser Val Leu	
35 40 45	
tta aat aaa cgt aca ata tca tta aca cca aca tca tct gac tcc att	192
Leu Asn Lys Arg Thr Ile Ser Leu Thr Pro Thr Ser Ser Asp Ser Ile	
50 55 60	
tat gat aga aat att atc acg aaa aag cca cac gaa atc aac tta tct	240
Tyr Asp Arg Asn Ile Ile Thr Lys Lys Pro His Glu Ile Asn Leu Ser	
65 70 75 80	
tcg tta tca ttt ttg ttt tgt gag att att agt tgg gca cac tct aat	288
Ser Leu Ser Phe Leu Phe Cys Glu Ile Ile Ser Trp Ala His Ser Asn	
85 90 95	
tcc aaa ggc att caa gat tta gaa aat cgt tta aac gga tta ggt tat	336
Ser Lys Gly Ile Gln Asp Leu Glu Asn Arg Leu Asn Gly Leu Gly Tyr	
100 105 110	
caa ata ggt caa cga tat ctc gaa ttg tgt aaa ata aga gaa ggt ttt	384
Gln Ile Gly Gln Arg Tyr Leu Glu Leu Cys Lys Ile Arg Glu Gly Phe	
115 120 125	
aaa aac agt aaa cga gag att aga ctt ttg gaa atg tta caa ttt att	432
Lys Asn Ser Lys Arg Glu Ile Arg Leu Leu Glu Met Leu Gln Phe Ile	
130 135 140	
cat ggt ccg ttc tgg aaa ttg att ttt ggt aaa act gct aat gaa tta	480
His Gly Pro Phe Trp Lys Leu Ile Phe Gly Lys Thr Ala Asn Glu Leu	
145 150 155 160	
gaa aaa tcg caa gat ttg ccc aat gaa tat atg att gtg gag aat gtg	528
Glu Lys Ser Gln Asp Leu Pro Asn Glu Tyr Met Ile Val Glu Asn Val	
165 170 175	

cca tta tta aat aga ttt att agt ata cct aag gag tat ggc gac tta 576  
 Pro Leu Leu Asn Arg Phe Ile Ser Ile Pro Lys Glu Tyr Gly Asp Leu  
 180 185 190

aat tgt tca gca ttt gtt gcg ggt ata att gag gga gca ctt gat aat 624  
 Asn Cys Ser Ala Phe Val Ala Gly Ile Ile Glu Gly Ala Leu Asp Asn  
 195 200 205

agt gga ttc aat gcc gat gtt aca gca cac acg gtc gct aca gat gca 672  
 Ser Gly Phe Asn Ala Asp Val Thr Ala His Thr Val Ala Thr Asp Ala  
 210 215 220

aat cca tta aga aca gta ttt ttg atc aag ttt gac gat tct gtt tta 720  
 Asn Pro Leu Arg Thr Val Phe Leu Ile Lys Phe Asp Asp Ser Val Leu  
 225 230 235 240

att aga gag agt ttg aga ttt gga taa 747  
 Ile Arg Glu Ser Leu Arg Phe Gly  
 245

<210> 2

<211> 249

<212>

<213> Candida albicans

<400> 2

Met Ser Asn Asp Asp Ile Ile Leu Pro Ser Val Ser Ser Leu Ser Lys  
 1 5 10 15

Leu Thr Ile Asn Asp Val Ser Lys Ser Gly Phe Gly Tyr Asn Pro Ser  
 20 25 30

Ile Gly Pro Ile Ser Asn Thr Ile Thr Leu Glu Ser Ser Ser Val Leu  
 35 40 45

Leu Asn Lys Arg Thr Ile Ser Leu Thr Pro Thr Ser Ser Asp Ser Ile  
 50 55 60

Tyr Asp Arg Asn Ile Ile Thr Lys Lys Pro His Glu Ile Asn Leu Ser  
 65 70 75 80

Ser Leu Ser Phe Leu Phe Cys Glu Ile Ile Ser Trp Ala His Ser Asn  
 85 90 95

Ser Lys Gly Ile Gln Asp Leu Glu Asn Arg Leu Asn Gly Leu Gly Tyr  
 100 105 110

Gln Ile Gly Gln Arg Tyr Leu Glu Leu Cys Lys Ile Arg Glu Gly Phe  
 115 120 125

Lys Asn Ser Lys Arg Glu Ile Arg Leu Leu Glu Met Leu Gln Phe Ile  
 130 135 140

His Gly Pro Phe Trp Lys Leu Ile Phe Gly Lys Thr Ala Asn Glu Leu  
 145 150 155 160

Glu Lys Ser Gln Asp Leu Pro Asn Glu Tyr Met Ile Val Glu Asn Val  
 165 170 175

Pro Leu Leu Asn Arg Phe Ile Ser Ile Pro Lys Glu Tyr Gly Asp Leu  
 180 185 190

Asn Cys Ser Ala Phe Val Ala Gly Ile Ile Glu Gly Ala Leu Asp Asn  
 195 200 205

Ser Gly Phe Asn Ala Asp Val Thr Ala His Thr Val Ala Thr Asp Ala  
 210 215 220

Asn Pro Leu Arg Thr Val Phe Leu Ile Lys Phe Asp Asp Ser Val Leu  
 225 230 235 240

Ile Arg Glu Ser Leu Arg Phe Gly  
 245



&lt;210&gt; 3

&lt;211&gt; 711

&lt;212&gt; ADN

&lt;213&gt; Candida albicans

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(711)

&lt;220&gt;

&lt;221&gt; modified\_base

&lt;222&gt; (577)..(579)

&lt;400&gt; 3

atg gat att gac gat att tta aaa gaa ttt gaa gag tct tca aaa gat 48  
 Met Asp Ile Asp Asp Ile Leu Lys Glu Phe Glu Glu Ser Ser Lys Asp

1 5 10 15

gaa aag att agc agt aaa aca tcg tct atc aac tta tat caa gac ttg 96  
 Glu Lys Ile Ser Ser Lys Thr Ser Ser Ile Asn Leu Tyr Gln Asp Leu

20 25 30

cta aga gct atg atc aac gaa cgt atg gct ccg gaa tta ttg cca tac 144  
 Leu Arg Ala Met Ile Asn Glu Arg Met Ala Pro Glu Leu Leu Pro Tyr

35 40 45

aaa caa gat tta atg tcc act gtt tta aca atg atg tct aac caa caa 192  
 Lys Gln Asp Leu Met Ser Thr Val Leu Thr Met Met Ser Asn Gln Gln

50 55 60

caa tat tta tta gaa tct cac gaa tat ggt gat atg aat ggc gac agt 240  
 Gln Tyr Leu Leu Glu Ser His Glu Tyr Gly Asp Met Asn Gly Asp Ser

65 70 75 80

ggt gta tta tcc gga gac ttt aaa tta caa cta atg att atc gaa act 288  
 Gly Val Leu Ser Gly Asp Phe Lys Leu Gln Leu Met Ile Ile Glu Thr

85 90 95

gat tta gag cgt ctc aac tat att gtt cga tta tac ata cga act cga	336
Asp Leu Glu Arg Leu Asn Tyr Ile Val Arg Leu Tyr Ile Arg Thr Arg	
100 105 110	
ttg agt aag ttg aat aaa ttt act att ttt tac atc aat gaa agc agt	384
Leu Ser Lys Leu Asn Lys Phe Thr Ile Phe Tyr Ile Asn Glu Ser Ser	
115 120 125	
caa aat gat aat tta ttg tcc aaa gag gaa aga gat tat ata cac aaa	432
Gln Asn Asp Asn Leu Leu Ser Lys Glu Glu Arg Asp Tyr Ile His Lys	
130 135 140	
tat ttc cag att ttg act caa tta tat aac aac tgt ttc ctc aaa aaa	480
Tyr Phe Gln Ile Leu Thr Gln Leu Tyr Asn Asn Cys Phe Leu Lys Lys	
145 150 155 160	
cta cca caa atg ttg acc tat ttg gat gac acc agt ggt gga caa tca	528
Leu Pro Gln Met Leu Thr Tyr Leu Asp Asp Thr Ser Gly Gly Gln Ser	
165 170 175	
atg atc gtt gag cca gat tta gac cag cct gtg ttt atc aaa tgt acc	576
Met Ile Val Glu Pro Asp Leu Asp Gln Pro Val Phe Ile Lys Cys Thr	
180 185 190	
ctg gaa gtc cca ata tta cta gat tac gac ggt gct aca gag ata gat	624
Ser Glu Val Pro Ile Leu Leu Asp Tyr Asp Gly Ala Thr Glu Ile Asp	
195 200 205	
tta gaa tta ata aaa aag gga gtc tac gtg gtg aaa tac agc cta gtc	672
Leu Glu Leu Ile Lys Lys Gly Val Tyr Val Val Lys Tyr Ser Leu Val	
210 215 220	
aaa aga tat att gat att gga gat gtg gta ttg ata tga	711
Lys Arg Tyr Ile Asp Ile Gly Asp Val Val Leu Ile	
225 230 235	

&lt;210&gt; 4

&lt;211&gt; 237

&lt;212&gt;

&lt;213&gt; Candida albicans

&lt;400&gt; 4

Met Asp Ile Asp Asp Ile Leu Lys Glu Phe Glu Glu Ser Ser Lys Asp

1

5

10

15

Glu Lys Ile Ser Ser Lys Thr Ser Ser Ile Asn Leu Tyr Gln Asp Leu

20

25

30

Leu Arg Ala Met Ile Asn Glu Arg Met Ala Pro Glu Leu Leu Pro Tyr

35

40

45

Lys Gln Asp Leu Met Ser Thr Val Leu Thr Met Met Ser Asn Gln Gln

50

55

60

Gln Tyr Leu Leu Glu Ser His Glu Tyr Gly Asp Met Asn Gly Asp Ser

65

70

75

80

Gly Val Leu Ser Gly Asp Phe Lys Leu Gln Leu Met Ile Ile Glu Thr

85

90

95

Asp Leu Glu Arg Leu Asn Tyr Ile Val Arg Leu Tyr Ile Arg Thr Arg

100

105

110

Leu Ser Lys Leu Asn Lys Phe Thr Ile Phe Tyr Ile Asn Glu Ser Ser

115

120

125

Gln Asn Asp Asn Leu Leu Ser Lys Glu Glu Arg Asp Tyr Ile His Lys

130

135

140

Tyr Phe Gln Ile Leu Thr Gln Leu Tyr Asn Asn Cys Phe Leu Lys Lys

145

150

155

160

Leu Pro Gln Met Leu Thr Tyr Leu Asp Asp Thr Ser Gly Gly Gln Ser

165

170

175

47

Met Ile Val Glu Pro Asp Leu Asp Gln Pro Val Phe Ile Lys Cys Thr  
180 185 190

Ser Glu Val Pro Ile Leu Leu Asp Tyr Asp Gly Ala Thr Glu Ile Asp  
195 200 205

Leu Glu Leu Ile Lys Lys Gly Val Tyr Val Val Lys Tyr Ser Leu Val  
210 215 220

Lys Arg Tyr Ile Asp Ile Gly Asp Val Val Leu Ile  
225 230 235

<210> 5

<211> 1383

<212> ADN

<213> Candida albicans

<220>

<221> CDS

<222> (1)..(1383)

<400> 5

atg gat ttc ata gga gag att ata gag cat gag aca gag gca cct aaa 48  
Met Asp Phe Ile Gly Glu Ile Ile Glu His Glu Thr Glu Ala Pro Lys  
1 5 10 15

gaa cca acc cca aaa ccc aca att ggt gga ttc ccc gaa ctt aaa aaa 96  
Glu Pro Thr Pro Lys Pro Thr Ile Gly Gly Phe Pro Glu Leu Lys Lys  
20 25 30

tta aaa gaa aag aaa gtc tca aga tgg agg caa aag caa caa cag gaa 144  
Leu Lys Glu Lys Lys Val Ser Arg Trp Arg Gln Lys Gln Gln Gln  
35 40 45

cag agc aca act tcc cca aaa act act gaa atc cgt tca gag gct tcc 192  
Gln Ser Thr Thr Ser Pro Lys Thr Thr Glu Ile Arg Ser Glu Ala Ser  
50 55 60

aaa att cac caa gaa aat atc gag aag atg gct caa atg tca gag gaa 240  
 Lys Ile His Gln Glu Asn Ile Glu Lys Met Ala Gln Met Ser Glu Glu  
 65 70 75 80

gag att ttg caa gag cgt gag gag tta cta aag ggt tta gat cct aaa 288  
 Glu Ile Leu Gln Glu Arg Glu Glu Leu Leu Lys Gly Leu Asp Pro Lys  
 85 90 95

tta att gaa agt ttg att ggt aga tcc aag aaa agg gaa gca aca gac 336  
 Leu Ile Glu Ser Leu Ile Gly Arg Ser Lys Lys Arg Glu Ala Thr Asp  
 100 105 110

cat gaa cac aat gga cat gct cat gaa cat gca gag gga tac cat gga 384  
 His Glu His Asn Gly His Ala His Glu His Ala Glu Gly Tyr His Gly  
 115 120 125

tgg att gga tca atg aaa act tct gaa gga tta aca gat tta tct caa 432  
 Trp Ile Gly Ser Met Lys Thr Ser Glu Gly Leu Thr Asp Leu Ser Gln  
 130 135 140

tta gat aag gaa gat gtg gac cgt gca ttg ggt ata agt tca tta tcc 480  
 Leu Asp Lys Glu Asp Val Asp Arg Ala Leu Gly Ile Ser Ser Leu Ser  
 145 150 155 160

tta tct gaa cct gag ggt ggc agt aat acg aaa aaa gtc gct ttc gac 528  
 Leu Ser Glu Pro Glu Gly Gly Ser Asn Thr Lys Lys Val Ala Phe Asp  
 165 170 175

gat aat atc aag acg gtt aaa ttt gaa gat ttg gat gat gga att gaa 576  
 Asp Asn Ile Lys Thr Val Lys Phe Glu Asp Leu Asp Asp Gly Ile Glu  
 180 185 190

ttg gat cca aat gga tgg gag gac gtt act gat gtc aat gaa tta gtt 624  
 Leu Asp Pro Asn Gly Trp Glu Asp Val Thr Asp Val Asn Glu Leu Val  
 195 200 205

cct aat aat gat cac att gca cct gac gat tac cag att aat cct gat 672

Pro Asn Asn Asp His Ile Ala Pro Asp Asp Tyr Gln Ile Asn Pro Asp  
 210 215 220

agc gat gaa gaa gga ttg aat aat act gtt cat ttt aca aaa ccc aaa 720  
 Ser Asp Glu Glu Gly Leu Asn Asn Thr Val His Phe Thr Lys Pro Lys  
 225 230 235 240

cag cca gat ttg gat ata aat gat ccc gat ttc ttt gat aag cta cat 768  
 Gln Pro Asp Leu Asp Ile Asn Asp Pro Asp Phe Phe Asp Lys Leu His  
 245 250 255

gag aaa tac tat cct gat ttg cct aaa gaa aca gaa aag ttg tca tgg 816  
 Glu Lys Tyr Tyr Pro Asp Leu Pro Lys Glu Thr Glu Lys Leu Ser Trp  
 260 265 270

atg aca cag cca atg cca aaa caa ttg tct acc gtt tat gaa tca ata 864  
 Met Thr Gln Pro Met Pro Lys Gln Leu Ser Thr Val Tyr Glu Ser Ile  
 275 280 285

tct gat atg aga ttt gac ttt aaa gga gat tta att gaa ttg ggt cca 912  
 Ser Asp Met Arg Phe Asp Phe Lys Gly Asp Leu Ile Glu Leu Gly Pro  
 290 295 300

gag gga gaa gaa cca aaa gat agt tca tcc gaa ata cct act tat atg 960  
 Glu Gly Glu Glu Pro Lys Asp Ser Ser Ser Glu Ile Pro Thr Tyr Met  
 305 310 315 320

gga ctt cat cat cat tcg gag aac cca cat atg gca ggt tat aca ttg 1008  
 Gly Leu His His His Ser Glu Asn Pro His Met Ala Gly Tyr Thr Leu  
 325 330 335

ggt gag ttg gca cat tta gcc aga tcg act tta gct gga caa aga tgc 1056  
 Gly Glu Leu Ala His Leu Ala Arg Ser Thr Leu Ala Gly Gln Arg Cys  
 340 345 350

ttg agc att caa aca tta ggg aga atc tta cat aaa ttg gga tta cat 1104  
 Leu Ser Ile Gln Thr Leu Gly Arg Ile Leu His Lys Leu Gly Leu His  
 355 360 365

aaa tac agt ata cta cca aaa aca gac tca gat gat cag agt ttt aca 1152  
 Lys Tyr Ser Ile Leu Pro Lys Thr Asp Ser Asp Asp Gln Ser Phe Thr  
 370 375 380

gat gaa atc aaa caa cta tca ctt gac ttt gaa gat atg atg tgg gac 1200  
 Asp Glu Ile Lys Gln Leu Ser Leu Asp Phe Glu Asp Met Met Trp Asp  
 385 390 395 400

ttg ata gac caa tta cga atc att gaa aca ata aca gag gca gct gat 1248  
 Leu Ile Asp Gln Leu Arg Ile Ile Glu Thr Ile Thr Glu Ala Ala Asp  
 405 410 415

gaa aaa aag acc aga aac tta tct gtc aga aat tat gca ata gag gca 1296  
 Glu Lys Lys Thr Arg Asn Leu Ser Val Arg Asn Tyr Ala Ile Glu Ala  
 420 425 430

ttg tgg tta tat aga act gga ggt gga aga cca gag ata act aaa caa 1344  
 Leu Trp Leu Tyr Arg Thr Gly Gly Gly Arg Pro Glu Ile Thr Lys Gln  
 435 440 445

acc gaa gag gat ttg ata gca caa gca gtt cag aaa taa 1383  
 Thr Glu Glu Asp Leu Ile Ala Gln Ala Val Gln Lys  
 450 455 460

<210> 6

<211> 461

<212>

<213> Candida albicans

<400> 6

Met Asp Phe Ile Gly Glu Ile Ile Glu His Glu Thr Glu Ala Pro Lys  
 1 5 10 15

Glu Pro Thr Pro Lys Pro Thr Ile Gly Gly Phe Pro Glu Leu Lys Lys  
 20 25 30

51

Leu Lys Glu Lys Lys Val Ser Arg Trp Arg Gln Lys Gln Gln Glu  
35 40 45

Gln Ser Thr Thr Ser Pro Lys Thr Thr Glu Ile Arg Ser Glu Ala Ser  
50 55 60

Lys Ile His Gln Glu Asn Ile Glu Lys Met Ala Gln Met Ser Glu Glu  
65 70 75 80

Glu Ile Leu Gln Glu Arg Glu Glu Leu Leu Lys Gly Leu Asp Pro Lys  
85 90 95

Leu Ile Glu Ser Leu Ile Gly Arg Ser Lys Lys Arg Glu Ala Thr Asp  
100 105 110

His Glu His Asn Gly His Ala His Glu His Ala Glu Gly Tyr His Gly  
115 120 125

Trp Ile Gly Ser Met Lys Thr Ser Glu Gly Leu Thr Asp Leu Ser Gln  
130 135 140

Leu Asp Lys Glu Asp Val Asp Arg Ala Leu Gly Ile Ser Ser Leu Ser  
145 150 155 160

Leu Ser Glu Pro Glu Gly Gly Ser Asn Thr Lys Lys Val Ala Phe Asp  
165 170 175

Asp Asn Ile Lys Thr Val Lys Phe Glu Asp Leu Asp Asp Gly Ile Glu  
180 185 190

Leu Asp Pro Asn Gly Trp Glu Asp Val Thr Asp Val Asn Glu Leu Val  
195 200 205

Pro Asn Asn Asp His Ile Ala Pro Asp Asp Tyr Gln Ile Asn Pro Asp  
210 215 220

Ser Asp Glu Glu Gly Leu Asn Asn Thr Val His Phe Thr Lys Pro Lys  
225 230 235 240



Gln Pro Asp Leu Asp Ile Asn Asp Pro Asp Phe Phe Asp Lys Leu His  
 245 250 255

Glu Lys Tyr Tyr Pro Asp Leu Pro Lys Glu Thr Glu Lys Leu Ser Trp  
 260 265 270

Met Thr Gln Pro Met Pro Lys Gln Leu Ser Thr Val Tyr Glu Ser Ile  
 275 280 285

Ser Asp Met Arg Phe Asp Phe Lys Gly Asp Leu Ile Glu Leu Gly Pro  
 290 295 300

Glu Gly Glu Glu Pro Lys Asp Ser Ser Ser Glu Ile Pro Thr Tyr Met  
 305 310 315 320

Gly Leu His His His Ser Glu Asn Pro His Met Ala Gly Tyr Thr Leu  
 325 330 335

Gly Glu Leu Ala His Leu Ala Arg Ser Thr Leu Ala Gly Gln Arg Cys  
 340 345 350

Leu Ser Ile Gln Thr Leu Gly Arg Ile Leu His Lys Leu Gly Leu His  
 355 360 365

Lys Tyr Ser Ile Leu Pro Lys Thr Asp Ser Asp Asp Gln Ser Phe Thr  
 370 375 380

Asp Glu Ile Lys Gln Leu Ser Leu Asp Phe Glu Asp Met Met Trp Asp  
 385 390 395 400

Leu Ile Asp Gln Leu Arg Ile Ile Glu Thr Ile Thr Glu Ala Ala Asp  
 405 410 415

Glu Lys Lys Thr Arg Asn Leu Ser Val Arg Asn Tyr Ala Ile Glu Ala  
 420 425 430

Leu Trp Leu Tyr Arg Thr Gly Gly Gly Arg Pro Glu Ile Thr Lys Gln  
 435 440 445

Thr Glu Glu Asp Leu Ile Ala Gln Ala Val Gln Lys

450

455

460

&lt;210&gt; 7

&lt;211&gt; 1383

&lt;212&gt; ADN

&lt;213&gt; Candida albicans

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (1380)

&lt;400&gt; 7

atg gat ttc ata gga gag att ata gag cat gag aca gag gca cct aaa 48

Met Asp Phe Ile Gly Glu Ile Ile Glu His Glu Thr Glu Ala Pro Lys

1

5

10

15

gaa cca acc cca aaa ccc aca att ggt gga ttc ccc gaa ctt aaa aaa 96

Glu Pro Thr Pro Lys Pro Thr Ile Gly Gly Phe Pro Glu Leu Lys Lys

20

25

30

tta aaa gaa aag aaa gtc tca aga tgg agg caa aag caa caa cag gag 144

Leu Lys Glu Lys Lys Val Ser Arg Trp Arg Gln Lys Gln Gln Gln Glu

35

40

45

cag agc aca act tcc cca aaa act act gaa atc cgt tca gag gct tcc 192

Gln Ser Thr Thr Ser Pro Lys Thr Thr Glu Ile Arg Ser Glu Ala Ser

50

55

60

aaa att cac caa gaa aat atc gag aag atg gct caa atg tca gag gaa 240

Lys Ile His Gln Glu Asn Ile Glu Lys Met Ala Gln Met Ser Glu Glu

65

70

75

80

gag att ttg caa gag cgt gag gag tta cta aag ggt tta gac cct aaa 288

Glu Ile Leu Gln Glu Arg Glu Glu Leu Leu Lys Gly Leu Asp Pro Lys

85

90

95

cat gaa cac aat gga cat gct cat gaa cat gca gag gga tac cat gga 384  
His Glu His Asn Gly His Ala His Glu His Ala Glu Gly Tyr His Gly  
115 120 125



agc gat gaa gaa gga ttg aat aat act gtt cat ttt aca aaa ccc aaa 720  
 Ser Asp Glu Glu Gly Leu Asn Asn Thr Val His Phe Thr Lys Pro Lys  
 225 230 235 240

cag cca gat ttg gat ata aat gat ccc gat ttc ttt gat aag cta cat	768
Gln Pro Asp Leu Asp Ile Asn Asp Pro Asp Phe Phe Asp Lys Leu His	
245 250 255	
 gag aaa tac tat cct gat ttg cct aaa gaa aca gaa aag ttg tca tgg	816
Glu Lys Tyr Tyr Pro Asp Leu Pro Lys Glu Thr Glu Lys Leu Ser Trp	
260 265 270	
 atg aca cag cca atg cca aaa caa ttg tct aca gtt tat gaa tca ata	864
Met Thr Gln Pro Met Pro Lys Gln Leu Ser Thr Val Tyr Glu Ser Ile	
275 280 285	
 tct gat atg aga ttt gac ttc aaa gga gat tta att gaa ttg agc gca	912
Ser Asp Met Arg Phe Asp Phe Lys Gly Asp Leu Ile Glu Leu Ser Ala	
290 295 300	
 gag gga gaa gaa cca aaa gat agt tca ttc gaa ata cct act tat atg	960
Glu Gly Glu Glu Pro Lys Asp Ser Ser Phe Glu Ile Pro Thr Tyr Met	
305 310 315 320	
 gga ctt cat cat cat tcg gag aac cca cat atg gca ggt tat aca ttg	1008
Gly Leu His His His Ser Glu Asn Pro His Met Ala Gly Tyr Thr Leu	
325 330 335	
 ggt gag ttg gca cat tta gcc aga tcg act tta gct gga caa aga tgc	1056
Gly Glu Leu Ala His Leu Ala Arg Ser Thr Leu Ala Gly Gln Arg Cys	
340 345 350	
 ttg agc att caa aca tta ggg aga ata tta cat aaa ttg gga tta cat	1104
Leu Ser Ile Gln Thr Leu Gly Arg Ile Leu His Lys Leu Gly Leu His	
355 360 365	
 aaa tac agt ata cta cca aaa aca gac tca gat gat cag agt ttt aca	1152
Lys Tyr Ser Ile Leu Pro Lys Thr Asp Ser Asp Asp Gln Ser Phe Thr	
370 375 380	

56

gat gaa atc aaa caa cta tca ctt gac ttt gaa gat atg atg tgg gac 1200  
 Asp Glu Ile Lys Gln Leu Ser Leu Asp Phe Glu Asp Met Met Trp Asp  
 385 390 395 400

ttg ata gac caa tta cga atc att gaa aca ata aca gag gca gct gat 1248  
 Leu Ile Asp Gln Leu Arg Ile Ile Glu Thr Ile Thr Glu Ala Ala Asp  
 405 410 415

gaa aaa aag acc aga aac tta tct gtc aga aat tat gca ata gag gca 1296  
 Glu Lys Lys Thr Arg Asn Leu Ser Val Arg Asn Tyr Ala Ile Glu Ala  
 420 425 430

ttg tgg tta tat aga act gga ggt gga aga cca gag ata act aaa caa 1344  
 Leu Trp Leu Tyr Arg Thr Gly Gly Gly Arg Pro Glu Ile Thr Lys Gln  
 435 440 445

acc gaa gag gat ttg ata gca caa gca gtt cag aaa taa 1383  
 Thr Glu Glu Asp Leu Ile Ala Gln Ala Val Gln Lys  
 450 455 460

<210> 8

<211> 460

<212>

<213> Candida albicans

<400> 8

Met Asp Phe Ile Gly Glu Ile Ile Glu His Glu Thr Glu Ala Pro Lys  
 1 5 10 15

Glu Pro Thr Pro Lys Pro Thr Ile Gly Gly Phe Pro Glu Leu Lys Lys  
 20 25 30

Leu Lys Glu Lys Lys Val Ser Arg Trp Arg Gln Lys Gln Gln Glu  
 35 40 45

Gln Ser Thr Thr Ser Pro Lys Thr Thr Glu Ile Arg Ser Glu Ala Ser  
 50 55 60

57

Lys	Ile	His	Gln	Glu	Asn	Ile	Glu	Lys	Met	Ala	Gln	Met	Ser	Glu	Glu
65					70				75					80	
Glu	Ile	Leu	Gln	Glu	Arg	Glu	Glu	Leu	Leu	Lys	Gly	Leu	Asp	Pro	Lys
			85					90					95		
Leu	Ile	Glu	Ser	Leu	Ile	Gly	Arg	Ser	Lys	Lys	Arg	Glu	Ala	Thr	Asp
		100					105					110			
His	Glu	His	Asn	Gly	His	Ala	His	Glu	His	Ala	Glu	Gly	Tyr	His	Gly
		115					120					125			
Trp	Ile	Gly	Ser	Met	Lys	Thr	Ser	Glu	Gly	Leu	Thr	Asp	Leu	Ser	Gln
		130				135					140				
Leu	Asp	Lys	Glu	Asp	Val	Asp	Arg	Ala	Leu	Gly	Ile	Ser	Ser	Leu	Ser
145				150					155					160	
Leu	Ser	Glu	Pro	Glu	Gly	Gly	Ser	Asn	Thr	Lys	Lys	Val	Ala	Phe	Asp
			165					170						175	
Asp	Asn	Ile	Lys	Thr	Val	Lys	Phe	Glu	Ala	Leu	Asp	Asp	Glu	Ile	Glu
		180						185					190		
Leu	Asp	Pro	Asn	Gly	Trp	Glu	Asp	Val	Thr	Asp	Val	Asn	Glu	Leu	Val
		195				200							205		
Pro	Asn	Asn	Asp	His	Ile	Ala	Pro	Asp	Asp	Tyr	Gln	Ile	Asn	Pro	Asp
		210				215						220			
Ser	Asp	Glu	Glu	Gly	Leu	Asn	Asn	Thr	Val	His	Phe	Thr	Lys	Pro	Lys
225				230						235				240	
Gln	Pro	Asp	Leu	Asp	Ile	Asn	Asp	Pro	Asp	Phe	Phe	Asp	Lys	Leu	His
			245					250					255		
Glu	Lys	Tyr	Tyr	Pro	Asp	Leu	Pro	Lys	Glu	Thr	Glu	Lys	Leu	Ser	Trp
		260						265					270		

Met Thr Gln Pro Met Pro Lys Gln Leu Ser Thr Val Tyr Glu Ser Ile  
 275 280 285

Ser Asp Met Arg Phe Asp Phe Lys Gly Asp Leu Ile Glu Leu Ser Ala  
 290 295 300

Glu Gly Glu Glu Pro Lys Asp Ser Ser Phe Glu Ile Pro Thr Tyr Met  
 305 310 315 320

Gly Leu His His His Ser Glu Asn Pro His Met Ala Gly Tyr Thr Leu  
 325 330 335

Gly Glu Leu Ala His Leu Ala Arg Ser Thr Leu Ala Gly Gln Arg Cys  
 340 345 350

Leu Ser Ile Gln Thr Leu Gly Arg Ile Leu His Lys Leu Gly Leu His  
 355 360 365

Lys Tyr Ser Ile Leu Pro Lys Thr Asp Ser Asp Asp Gln Ser Phe Thr  
 370 375 380

Asp Glu Ile Lys Gln Leu Ser Leu Asp Phe Glu Asp Met Met Trp Asp  
 385 390 395 400

Leu Ile Asp Gln Leu Arg Ile Ile Glu Thr Ile Thr Glu Ala Ala Asp  
 405 410 415

Glu Lys Lys Thr Arg Asn Leu Ser Val Arg Asn Tyr Ala Ile Glu Ala  
 420 425 430

Leu Trp Leu Tyr Arg Thr Gly Gly Gly Arg Pro Glu Ile Thr Lys Gln  
 435 440 445

Thr Glu Glu Asp Leu Ile Ala Gln Ala Val Gln Lys  
 450 455 460

<210> 9  
 <211> 2262  
 <212> ADN  
 <213> Candida albicans

<220>  
 <221> CDS  
 <222> (1)..(2262)

<220>  
 <221> modified\_base  
 <222> (1093)..(1095)

<220>  
 <221> modified\_base  
 <222> (1828)..(1830)

<400> 9  
 atg gca gca gca cca cca cca cca gcg aaa aac cag ggt aag gca aaa 48  
 Met Ala Ala Ala Pro Pro Pro Pro Ala Lys Asn Gln Gly Lys Ala Lys  
 1 5 10 15  
 cag cat gtt aca ggt gcc agg ttc cgt cag cga aaa atc tcg gta aag 96  
 Gln His Val Thr Gly Ala Arg Phe Arg Gln Arg Lys Ile Ser Val Lys  
 20 25 30  
 cag ccc ttg act att tat aaa cag aga gac cta cct act cta gat agc 144  
 Gln Pro Leu Thr Ile Tyr Lys Gln Arg Asp Leu Pro Thr Leu Asp Ser  
 35 40 45  
 aat gag tta gag cct agt caa gtc cat cat tta aat tct aat gcg tca 192  
 Asn Glu Leu Glu Pro Ser Gln Val His His Leu Asn Ser Asn Ala Ser  
 50 55 60  
 tca tca tca aca caa caa ccg aga gac ctt cat gca gtt gaa act ggg 240  
 Ser Ser Ser Thr Gln Gln Pro Arg Asp Leu His Ala Val Glu Thr Gly  
 65 70 75 80



60

gtt gac aag aat gag gaa gag gaa gtg cat ctt cag caa gtt atc aat 288  
Val Asp Lys Asn Glu Glu Glu Glu Val His Leu Gln Gln Val Ile Asn

85

90

95

gct gca caa aaa gca ctt ttg ggt tcg aaa aaa gaa gaa aaa agc agt 336  
Ala Ala Gln Lys Ala Leu Leu Gly Ser Lys Lys Glu Glu Lys Ser Ser

100

105

110

gat atg tat att ccc aca ccg gac gct tcg agg ata tgg ccc gag gca 384  
Asp Met Tyr Ile Pro Thr Pro Asp Ala Ser Arg Ile Trp Pro Glu Ala

115

120

125

cac aag tat tac aag gat caa aag ttc aag cag cca gag aca tat atc 432  
His Lys Tyr Tyr Lys Asp Gln Lys Phe Lys Gln Pro Glu Thr Tyr Ile

130

135

140

aag ttt agt gcg aca gta gag gac aca gtg ggt gtg gag tac aat atg 480  
Lys Phe Ser Ala Thr Val Glu Asp Thr Val Gly Val Glu Tyr Asn Met

145

150

155

160

gac gag gta gat gaa aag ttt tat aga gag aca cta tgc aag tac tat 528  
Asp Glu Val Asp Glu Lys Phe Tyr Arg Glu Thr Leu Cys Lys Tyr Tyr

165

170

175

ccc aaa aag aaa aac aag tca gat gag aac aat cga aag tgt act gaa 576  
Pro Lys Lys Lys Asn Lys Ser Asp Glu Asn Asn Arg Lys Cys Thr Glu

180

185

190

ttg gag ttt gaa aca atc tgt gac aag ttg gaa aag acc att gaa gca 624  
Leu Glu Phe Glu Thr Ile Cys Asp Lys Leu Glu Lys Thr Ile Glu Ala

195

200

205

cga caa ccg ttt ttg tct atg gac ccc agc aac att cta tcg tac gag 672  
Arg Gln Pro Phe Leu Ser Met Asp Pro Ser Asn Ile Leu Ser Tyr Glu

210

215

220

61

gag ttg tcg tcg tac att gtg gat cag ttt aaa agt gca gtg aaa aca 720

Glu Leu Ser Ser Tyr Ile Val Asp Gln Phe Lys Ser Ala Val Lys Thr

225 230 235 240

agc aac ccg tat att gtt acc aat ggt ggg aat cta gag tat ata tcg 768

Ser Asn Pro Tyr Ile Val Thr Asn Gly Gly Asn Leu Glu Tyr Ile Ser

245 250 255

acg aca gct tta aaa gag aga ttg tcg aag gaa ata aag tat gaa ccg 816

Thr Thr Ala Leu Lys Glu Arg Leu Ser Lys Glu Ile Lys Tyr Glu Pro

260 265 270

ttt gtt act att ttt gat aag aac caa atg tcc aca agt gcg gtg aga 864

Phe Val Thr Ile Phe Asp Lys Asn Gln Met Ser Thr Ser Ala Val Arg

275 280 285

cct att ccc aaa ttg ttt gag ttg ttc ggc aga cct gtt tat gat cat 912

Pro Ile Pro Lys Leu Phe Glu Leu Phe Gly Arg Pro Val Tyr Asp His

290 295 300

tgg aag gag aga aaa ata gaa aga aag ggc aaa acc atc cag ccc aca 960

Trp Lys Glu Arg Lys Ile Glu Arg Lys Gly Lys Thr Ile Gln Pro Thr

305 310 315 320

ctc aaa ttt gag gat cct aac tcg aac gaa aag gaa aac gac aat gac 1008

Leu Lys Phe Glu Asp Pro Asn Ser Asn Glu Lys Glu Asn Asp Asn Asp

325 330 335

cca tat ata tgt ttc aga cga cgt gag ttt agg caa gca aga aag acg 1056

Pro Tyr Ile Cys Phe Arg Arg Arg Glu Phe Arg Gln Ala Arg Lys Thr

340 345 350

aga aga gcc gat aca att ggt gca gag aga ata aga ctg atg caa aag 1104

Arg Arg Ala Asp Thr Ile Gly Ala Glu Arg Ile Arg Ser Met Gln Lys

355 360 365

tcg ttg cac cgc gca cgt gat ttg ata atg agt gtt agt gaa aga gag 1152  
 Ser Leu His Arg Ala Arg Asp Leu Ile Met Ser Val Ser Glu Arg Glu

370

375

380

atc ctc aaa ctc gac aat ttt caa gca gag cat gaa ttg ttt aaa gcc 1200  
 Ile Leu Lys Leu Asp Asn Phe Gln Ala Glu His Glu Leu Phe Lys Ala

385

390

395

400

agg tgc gct acc aag gct tgt aag agg gag ctc aat atc aag ggt gac 1248  
 Arg Cys Ala Thr Lys Ala Cys Lys Arg Glu Leu Asn Ile Lys Gly Asp

405

410

415

gaa tac ttg ttc ttt ccg cat aaa aag aag aaa att gtt cgt act gaa 1296  
 Glu Tyr Leu Phe Phe Pro His Lys Lys Lys Lys Ile Val Arg Thr Glu

420

425

430

gat gaa gaa agg gag aag aag aga gaa aag aag aag caa gac caa gaa 1344  
 Asp Glu Glu Arg Glu Lys Lys Arg Glu Lys Lys Lys Gln Asp Gln Glu

435

440

445

ctt gca ctc aag caa caa caa gca cta cag caa cag cag caa caa cca 1392  
 Leu Ala Leu Lys Gln Gln Gln Ala Leu Gln Gln Gln Gln Gln Pro

450

455

460

cca caa cca cca caa caa gca cca tca aaa caa gat ggt aca tca acg 1440  
 Pro Gln Pro Pro Gln Gln Ala Pro Ser Lys Gln Asp Gly Thr Ser Thr

465

470

475

480

agc cag cct tat gtc aaa ctc cca ccc gca aaa gtt cca gat atg gat 1488  
 Ser Gln Pro Tyr Val Lys Leu Pro Pro Ala Lys Val Pro Asp Met Asp

485

490

495

ctt gtt aca gtt tcg ttg gta tta aag gaa aag aac gaa acc atc aaa 1536  
 Leu Val Thr Val Ser Leu Val Leu Lys Glu Lys Asn Glu Thr Ile Lys

500

505

510

63

cgt gct gtg ttg gag aaa ttg cgc aag aga aag gaa cac gac aag gga 1584  
Arg Ala Val Leu Glu Lys Leu Arg Lys Arg Lys Glu His Asp Lys Gly

515

520

525

ttt atc aat ttg aca gac gat ccg tat cag cca ttt ttc gat att tca 1632  
Phe Ile Asn Leu Thr Asp Asp Pro Tyr Gln Pro Phe Phe Asp Ile Ser

530

535

540

acc aat agg gcc gaa gag ttg agc cat att ccg tat tcg tcg att gcg 1680  
Thr Asn Arg Ala Glu Glu Leu Ser His Ile Pro Tyr Ser Ser Ile Ala

545

550

555

560

gcc aca cac tat cac caa ttc aac aca tcg aac tac atg aac gac caa 1728  
Ala Thr His Tyr His Gln Phe Asn Thr Ser Asn Tyr Met Asn Asp Gln

565

570

575

ctt aaa aag cta ctt gaa gag aaa aaa cct tta cct ggt gta aaa acg 1776  
Leu Lys Lys Leu Leu Glu Glu Lys Lys Pro Leu Pro Gly Val Lys Thr

580

585

590

ttt ttg ggt tct aac ggg gag ttg gta cca tcg aag gca ttt cca cat 1824  
Phe Leu Gly Ser Asn Gly Glu Leu Val Pro Ser Lys Ala Phe Pro His

595

600

605

ttg ctg tcg ttg ctt gag gaa aag tat aag gcg aca agt ggg tat att 1872  
Leu Ser Ser Leu Leu Glu Glu Lys Tyr Lys Ala Thr Ser Gly Tyr Ile

610

615

620

gaa cga tta ttg caa agc gtg gag acg caa gat ttt agt tca tac acc 1920  
Glu Arg Leu Leu Gln Ser Val Glu Thr Gln Asp Phe Ser Ser Tyr Thr

625

630

635

640

aat ggc ttt aaa gat gtt gag cca aaa gaa aca aat gaa cct gtt atg 1968  
Asn Gly Phe Lys Asp Val Glu Pro Lys Glu Thr Asn Glu Pro Val Met

645

650

655

64

gcg ttt ccc cag aga ata cgt cga aga gtg ggc agg gct ggc agg gtt 2016  
Ala Phe Pro Gln Arg Ile Arg Arg Arg Val Gly Arg Ala Gly Arg Val  
660 665 670

ttt ttg gac cac cag caa gag tac ccg caa ccg aat ttt cag caa gac 2064  
Phe Leu Asp His Gln Gln Glu Tyr Pro Gln Pro Asn Phe Gln Gln Asp  
675 680 685

aca gat cgt gtg gga ggt atc cca gat gtg tat tgt aaa gag gat gcc 2112  
Thr Asp Arg Val Gly Gly Ile Pro Asp Val Tyr Cys Lys Glu Asp Ala  
690 695 700

att aaa cga tta cag tca aag tgg aag ttc gat aca gaa tat aaa aca 2160  
Ile Lys Arg Leu Gln Ser Lys Trp Lys Phe Asp Thr Glu Tyr Lys Thr  
705 710 715 720

act gaa cca ttt agt ttg gat cct tca aag ttg aat ggt att agt cca 2208  
Thr Glu Pro Phe Ser Leu Asp Pro Ser Lys Leu Asn Gly Ile Ser Pro  
725 730 735

tct acg caa tcg att aga ttt ggg tct atg ttg ttg aat aga aca cgt 2256  
Ser Thr Gln Ser Ile Arg Phe Gly Ser Met Leu Leu Asn Arg Thr Arg  
740 745 750

aaa tag 2262  
Lys

<210> 10

<211> 754

<212>

<213> Candida albicans

<400> 10

Met Ala Ala Ala Pro Pro Pro Pro Ala Lys Asn Gln Gly Lys Ala Lys  
1 5 10 15

65

Gln His Val Thr Gly Ala Arg Phe Arg Gln Arg Lys Ile Ser Val Lys  
20 25 30

Gln Pro Leu Thr Ile Tyr Lys Gln Arg Asp Leu Pro Thr Leu Asp Ser  
35 40 45

Asn Glu Leu Glu Pro Ser Gln Val His His Leu Asn Ser Asn Ala Ser  
50 55 60

Ser Ser Ser Thr Gln Gln Pro Arg Asp Leu His Ala Val Glu Thr Gly  
65 70 75 80

Val Asp Lys Asn Glu Glu Glu Glu Val His Leu Gln Gln Val Ile Asn  
85 90 95

Ala Ala Gln Lys Ala Leu Leu Gly Ser Lys Lys Glu Glu Lys Ser Ser  
100 105 110

Asp Met Tyr Ile Pro Thr Pro Asp Ala Ser Arg Ile Trp Pro Glu Ala  
115 120 125

His Lys Tyr Tyr Lys Asp Gln Lys Phe Lys Gln Pro Glu Thr Tyr Ile  
130 135 140

Lys Phe Ser Ala Thr Val Glu Asp Thr Val Gly Val Glu Tyr Asn Met  
145 150 155 160

Asp Glu Val Asp Glu Lys Phe Tyr Arg Glu Thr Leu Cys Lys Tyr Tyr  
165 170 175

Pro Lys Lys Lys Asn Lys Ser Asp Glu Asn Asn Arg Lys Cys Thr Glu  
180 185 190

Leu Glu Phe Glu Thr Ile Cys Asp Lys Leu Glu Lys Thr Ile Glu Ala  
195 200 205

Arg Gln Pro Phe Leu Ser Met Asp Pro Ser Asn Ile Leu Ser Tyr Glu  
210 215 220

Glu Leu Ser Ser Tyr Ile Val Asp Gln Phe Lys Ser Ala Val Lys Thr  
 225                      230                      235                      240

Ser Asn Pro Tyr Ile Val Thr Asn Gly Gly Asn Leu Glu Tyr Ile Ser  
                          245                      250                      255

Thr Thr Ala Leu Lys Glu Arg Leu Ser Lys Glu Ile Lys Tyr Glu Pro  
                          260                      265                      270

Phe Val Thr Ile Phe Asp Lys Asn Gln Met Ser Thr Ser Ala Val Arg  
                          275                      280                      285

Pro Ile Pro Lys Leu Phe Glu Leu Phe Gly Arg Pro Val Tyr Asp His  
                          290                      295                      300

Trp Lys Glu Arg Lys Ile Glu Arg Lys Gly Lys Thr Ile Gln Pro Thr  
 305                      310                      315                      320

Leu Lys Phe Glu Asp Pro Asn Ser Asn Glu Lys Glu Asn Asp Asn Asp  
                          325                      330                      335

Pro Tyr Ile Cys Phe Arg Arg Arg Glu Phe Arg Gln Ala Arg Lys Thr  
                          340                      345                      350

Arg Arg Ala Asp Thr Ile Gly Ala Glu Arg Ile Arg Ser Met Gln Lys  
                          355                      360                      365

Ser Leu His Arg Ala Arg Asp Leu Ile Met Ser Val Ser Glu Arg Glu  
                          370                      375                      380

Ile Leu Lys Leu Asp Asn Phe Gln Ala Glu His Glu Leu Phe Lys Ala  
 385                      390                      395                      400

Arg Cys Ala Thr Lys Ala Cys Lys Arg Glu Leu Asn Ile Lys Gly Asp  
                          405                      410                      415

Glu Tyr Leu Phe Phe Pro His Lys Lys Lys Lys Ile Val Arg Thr Glu

420

425

430

Asp Glu Glu Arg Glu Lys Lys Arg Glu Lys Lys Lys Gln Asp Gln Glu

435

440

445

Leu Ala Leu Lys Gln Gln Gln Ala Leu Gln Gln Gln Gln Gln Pro

450

455

460

Pro Gln Pro Pro Gln Gln Ala Pro Ser Lys Gln Asp Gly Thr Ser Thr

465

470

475

480

Ser Gln Pro Tyr Val Lys Leu Pro Pro Ala Lys Val Pro Asp Met Asp

485

490

495

Leu Val Thr Val Ser Leu Val Leu Lys Glu Lys Asn Glu Thr Ile Lys

500

505

510

Arg Ala Val Leu Glu Lys Leu Arg Lys Arg Lys Glu His Asp Lys Gly

515

520

525

Phe Ile Asn Leu Thr Asp Asp Pro Tyr Gln Pro Phe Phe Asp Ile Ser

530

535

540

Thr Asn Arg Ala Glu Glu Leu Ser His Ile Pro Tyr Ser Ser Ile Ala

545

550

555

560

Ala Thr His Tyr His Gln Phe Asn Thr Ser Asn Tyr Met Asn Asp Gln

565

570

575

Leu Lys Lys Leu Leu Glu Glu Lys Lys Pro Leu Pro Gly Val Lys Thr

580

585

590

Phe Leu Gly Ser Asn Gly Glu Leu Val Pro Ser Lys Ala Phe Pro His

595

600

605

Leu Ser Ser Leu Leu Glu Glu Lys Tyr Lys Ala Thr Ser Gly Tyr Ile

610

615

620



Glu Arg Leu Leu Gln Ser Val Glu Thr Gln Asp Phe Ser Ser Tyr Thr  
 625                      630                      635                      640

Asn Gly Phe Lys Asp Val Glu Pro Lys Glu Thr Asn Glu Pro Val Met  
                     645                      650                      655

Ala Phe Pro Gln Arg Ile Arg Arg Arg Val Gly Arg Ala Gly Arg Val  
                     660                      665                      670

Phe Leu Asp His Gln Gln Glu Tyr Pro Gln Pro Asn Phe Gln Gln Asp  
                     675                      680                      685

Thr Asp Arg Val Gly Gly Ile Pro Asp Val Tyr Cys Lys Glu Asp Ala  
                     690                      695                      700

Ile Lys Arg Leu Gln Ser Lys Trp Lys Phe Asp Thr Glu Tyr Lys Thr  
 705                      710                      715                      720

Thr Glu Pro Phe Ser Leu Asp Pro Ser Lys Leu Asn Gly Ile Ser Pro  
                     725                      730                      735

Ser Thr Gln Ser Ile Arg Phe Gly Ser Met Leu Leu Asn Arg Thr Arg  
                     740                      745                      750

Lys

<210> 11

<211> 447

<212> ADN

<213> Candida albicans

<220>

<221> CDS

<222> (1)..(447)

&lt;400&gt; 11

atg tca gat ata gat ata gat aat gta tta aat tta gaa gaa gaa caa	48
Met Ser Asp Ile Asp Ile Asp Asn Val Leu Asn Leu Glu Glu Glu Gln	
1 5 10 15	
tat gaa tta gga ttt aaa gaa ggt caa ata caa gga aca aaa gat caa	96
Tyr Glu Leu Gly Phe Lys Glu Gly Gln Ile Gln Gly Thr Lys Asp Gln	
20 25 30	
tat tta gaa gga aaa gaa tat ggt tat caa act gga ttt caa cga ttt	144
Tyr Leu Glu Gly Lys Glu Tyr Gly Tyr Gln Thr Gly Phe Gln Arg Phe	
35 40 45	
tta atc att ggt tat att caa gaa tta atg aaa ttt tgg tta tcc cat	192
Leu Ile Ile Gly Tyr Ile Gln Glu Leu Met Lys Phe Trp Leu Ser His	
50 55 60	
ata gat caa tat aat aac tct tct tca ctt cgg aat cat ttg aat aat	240
Ile Asp Gln Tyr Asn Asn Ser Ser Ser Leu Arg Asn His Leu Asn Asn	
65 70 75 80	
ttg gaa gat att atg gca caa att tct ata acg aat gga gat aaa gaa	288
Leu Glu Asp Ile Met Ala Gln Ile Ser Ile Thr Asn Gly Asp Lys Glu	
85 90 95	
gtt gaa gat tat gaa aaa aat att aaa aag gca aga aat aaa tta aga	336
Val Glu Asp Tyr Glu Lys Asn Ile Lys Lys Ala Arg Asn Lys Leu Arg	
100 105 110	
gtg ata gct agt ata act aaa gaa act tgg aaa att gat tca ttg gat	384
Val Ile Ala Ser Ile Thr Lys Glu Thr Trp Lys Ile Asp Ser Leu Asp	
115 120 125	
aat ttg gtg aaa gaa gta ggt gga act tta caa gtt agt gaa aac ccc	432
Asn Leu Val Lys Glu Val Gly Gly Thr Leu Gln Val Ser Glu Asn Pro	
130 135 140	

gat gat atg tgg tga

Asp Asp Met Trp

145

<210> 12

<211> 149

<212>

<213> Candida albicans

<400> 12

Met Ser Asp Ile Asp Ile Asp Asn Val Leu Asn Leu Glu Glu Glu Gln

1

5

10

15

Tyr Glu Leu Gly Phe Lys Glu Gly Gln Ile Gln Gly Thr Lys Asp Gln

20

25

30

Tyr Leu Glu Gly Lys Glu Tyr Gly Tyr Gln Thr Gly Phe Gln Arg Phe

35

40

45

Leu Ile Ile Gly Tyr Ile Gln Glu Leu Met Lys Phe Trp Leu Ser His

50

55

60

Ile Asp Gln Tyr Asn Asn Ser Ser Ser Leu Arg Asn His Leu Asn Asn

65

70

75

80

Leu Glu Asp Ile Met Ala Gln Ile Ser Ile Thr Asn Gly Asp Lys Glu

85

90

95

Val Glu Asp Tyr Glu Lys Asn Ile Lys Lys Ala Arg Asn Lys Leu Arg

100

105

110

Val Ile Ala Ser Ile Thr Lys Glu Thr Trp Lys Ile Asp Ser Leu Asp

115

120

125

Asn Leu Val Lys Glu Val Gly Gly Thr Leu Gln Val Ser Glu Asn Pro

130

135

140

Asp Asp Met Trp

145

<210> 13

<211> 966

<212> ADN

<213> Candida albicans

<220>

<221> CDS

<222> (1) .. (966)

<400> 13

atg ggt aaa aga aga gta gat gaa gaa tct gat tca gat att gat gtt 48

Met Gly Lys Arg Arg Val Asp Glu Glu Ser Asp Ser Asp Ile Asp Val

1

5

10

15

agt tca acc gat tca gaa act gaa tta gaa agc aca caa caa caa caa 96

Ser Ser Thr Asp Ser Glu Thr Glu Leu Glu Ser Thr Gln Gln Gln Gln

20

25

30

caa caa caa gaa ggt gct act aca att caa gaa act gtt gat gtt gat 144

Gln Gln Gln Glu Gly Ala Thr Thr Ile Gln Glu Thr Val Asp Val Asp

35

40

45

ttt gat ttt ttt gat tta aat cct caa att gat ttc cat gct act aag 192

Phe Asp Phe Phe Asp Leu Asn Pro Gln Ile Asp Phe His Ala Thr Lys

50

55

60

aat ttt tta aga caa tta ttt ggt gat gat aat gga gaa ttt aat tta 240

Asn Phe Leu Arg Gln Leu Phe Gly Asp Asp Asn Gly Glu Phe Asn Leu

65

70

75

80

agt gaa ata gcc gat tta att tta cga gaa aat tcc gtg ggg aca tca 288

Ser Glu Ile Ala Asp Leu Ile Leu Arg Glu Asn Ser Val Gly Thr Ser

85

90

95

att aaa act gaa gga atg gaa agt gat cca ttt gca att tta agt gta 336  
 Ile Lys Thr Glu Gly Met Glu Ser Asp Pro Phe Ala Ile Leu Ser Val

100

105

110

att aat tta act aat aat tta aat gtg gcc gtg att aaa caa ttg att 384  
 Ile Asn Leu Thr Asn Asn Leu Asn Val Ala Val Ile Lys Gln Leu Ile

115

120

125

gaa tat att tca aat aaa acc aaa tct aaa act gaa ttc aat att att 432  
 Glu Tyr Ile Ser Asn Lys Thr Lys Ser Lys Thr Glu Phe Asn Ile Ile

130

135

140

ttg aaa aaa ttg tta acc aat cag aac gat act act aga gat agg aaa 480  
 Leu Lys Lys Leu Leu Thr Asn Gln Asn Asp Thr Thr Arg Asp Arg Lys

145

150

155

160

ttt aaa act gga tta ata att agt gaa aga ttt ata aat atg cca gtt 528  
 Phe Lys Thr Gly Leu Ile Ile Ser Glu Arg Phe Ile Asn Met Pro Val

165

170

175

gaa gtg att cca cca atg tat aaa atg ctt tta caa gaa atg gaa aaa 576  
 Glu Val Ile Pro Pro Met Tyr Lys Met Leu Leu Gln Glu Met Glu Lys

180

185

190

gct gaa gat gct cat gaa aat tat gaa ttt gat tat ttt tta att ata 624  
 Ala Glu Asp Ala His Glu Asn Tyr Glu Phe Asp Tyr Phe Leu Ile Ile

195

200

205

tca aga gtt tat caa tta gtt gat cca gtg gaa aga gaa gat gaa gat 672  
 Ser Arg Val Tyr Gln Leu Val Asp Pro Val Glu Arg Glu Asp Glu Asp

210

215

220

cac gaa aaa gaa tcc aat cgt aaa aag aag aac aag aat aag aag aag 720  
 His Glu Lys Glu Ser Asn Arg Lys Lys Lys Asn Lys Asn Lys Lys Lys

225

230

235

240

73

aaa ttg gct aat aat gaa cca aaa cca ata gaa atg gat tat ttc cat 768  
 Lys Leu Ala Asn Asn Glu Pro Lys Pro Ile Glu Met Asp Tyr Phe His  
 245 250 255

ctt gaa gat caa att ttg gaa tca aat act caa ttt aaa gga ata ttt 816  
 Leu Glu Asp Gln Ile Leu Glu Ser Asn Thr Gln Phe Lys Gly Ile Phe  
 260 265 270

gaa tat aat aat gaa aat aaa caa gaa aca gat tca aga aga gta ttt 864  
 Glu Tyr Asn Asn Glu Asn Lys Gln Glu Thr Asp Ser Arg Arg Val Phe  
 275 280 285

act gaa tat ggt att gat cct aaa tta agt tta atc tta att gat aaa 912  
 Thr Glu Tyr Gly Ile Asp Pro Lys Leu Ser Leu Ile Leu Ile Asp Lys  
 290 295 300

gat aat tta gct aaa tca gtc att gaa atg gaa caa caa ttc cca cct 960  
 Asp Asn Leu Ala Lys Ser Val Ile Glu Met Glu Gln Gln Phe Pro Pro  
 305 310 315 320

cca taa 966  
 Pro

<210> 14

<211> 322

<212>

<213> Candida albicans

<400> 14

Met Gly Lys Arg Arg Val Asp Glu Glu Ser Asp Ser Asp Ile Asp Val  
 1 5 10 15

Ser Ser Thr Asp Ser Glu Thr Glu Leu Glu Ser Thr Gln Gln Gln Gln  
 20 25 30

Gln Gln Gln Glu Gly Ala Thr Thr Ile Gln Glu Thr Val Asp Val Asp  
 35 40 45

Phe Asp Phe Phe Asp Leu Asn Pro Gln Ile Asp Phe His Ala Thr Lys  
 50 55 60

Asn Phe Leu Arg Gln Leu Phe Gly Asp Asp Asn Gly Glu Phe Asn Leu  
 65 70 75 80

Ser Glu Ile Ala Asp Leu Ile Leu Arg Glu Asn Ser Val Gly Thr Ser  
 85 90 95

Ile Lys Thr Glu Gly Met Glu Ser Asp Pro Phe Ala Ile Leu Ser Val  
 100 105 110

Ile Asn Leu Thr Asn Asn Leu Asn Val Ala Val Ile Lys Gln Leu Ile  
 115 120 125

Glu Tyr Ile Ser Asn Lys Thr Lys Ser Lys Thr Glu Phe Asn Ile Ile  
 130 135 140

Leu Lys Lys Leu Leu Thr Asn Gln Asn Asp Thr Thr Arg Asp Arg Lys  
 145 150 155 160

Phe Lys Thr Gly Leu Ile Ile Ser Glu Arg Phe Ile Asn Met Pro Val  
 165 170 175

Glu Val Ile Pro Pro Met Tyr Lys Met Leu Leu Gln Glu Met Glu Lys  
 180 185 190

Ala Glu Asp Ala His Glu Asn Tyr Glu Phe Asp Tyr Phe Leu Ile Ile  
 195 200 205

Ser Arg Val Tyr Gln Leu Val Asp Pro Val Glu Arg Glu Asp Glu Asp  
 210 215 220

His Glu Lys Glu Ser Asn Arg Lys Lys Lys Asn Lys Asn Lys Lys Lys  
 225 230 235 240

75

Lys Leu Ala Asn Asn Glu Pro Lys Pro Ile Glu Met Asp Tyr Phe His  
245 250 255

Leu Glu Asp Gln Ile Leu Glu Ser Asn Thr Gln Phe Lys Gly Ile Phe  
260 265 270

Glu Tyr Asn Asn Glu Asn Lys Gln Glu Thr Asp Ser Arg Arg Val Phe  
275 280 285

Thr Glu Tyr Gly Ile Asp Pro Lys Leu Ser Leu Ile Leu Ile Asp Lys  
290 295 300

Asp Asn Leu Ala Lys Ser Val Ile Glu Met Glu Gln Gln Phe Pro Pro  
305 310 315 320

Pro

<210> 15

<211> 320

<212> ADN

<213> Candida albicans

<400> 15

caatttattc atggtccggt ctggaaattg atttttggta aaactgctaa tgaattagaa 60

aaatcgcaag atttgcccaa tgaatatatg attgtggaga atgtgccatt attaaataga 120

tttattagta tacctaagga gtatggcgac ttaaattggt cagcatttgt tgcgggtata 180

attgagggag cacttgataa tagtggaattc aatgccgatg ttacagcaca cacggtcgct 240

acagatgcaa atccattaag aacagtattt ttgatcaagt ttgacgattc tgttttaatt 300

agagagagtt tgagatttgg 320



<210> 16

<211> 295

<212> ADN

<213> *Candida albicans*

<400> 16

gttcattgtt ggtgactcag agcgtctcaa ctatattgtt cgattataca tacgaactcg 60  
attgagtaag ttgaataaat ttactatatt ttacatcaat gaaagcagtc aaaatgataa 120  
tttattgtcc aaagaggaaa gagattatat acacaaatat ttccagattt tgactcaatt 180  
atataacaac tgtttcctca aaaaactacc acaaattgtt acctatttgg atgacaccag 240  
tggtggacaa tcaatgatcg ttgagccaga tttagaccag cctgtgttta tcaaa 295

<210> 17

<211> 392

<212> ADN

<213> *Candida albicans*

<400> 17

atctctgata tgagatttgg ctttaaaggc gatttaattg aattggctcc agtgggagat 60  
gcacaaaaag atagttcatc cgacatacgt actcatatgg gactccatca tcattcggag 120  
acccacata tggcagggtta tacattgggt gagttggccc atttagccag atcgacttta 180  
gctggacaaa gatgcttgag cattcaaaca ttagggagaa tcttcataa attgggatta 240  
cataaatata gtatactacc aaaccagctc aatgatcaga gttttacaga tgaatcaaaa 300  
ctatcacttg actttgaaga tagatgtggg acttgataga ccaattacga atcattgaaa 360  
caataacaga ggcagctgat ggaaaaaaga cc 392

<210> 18  
 <211> 335  
 <212> ADN  
 <213> Candida albicans

<400> 18  
 attccacac cggacgcttc gaggatatgg cccgaggcac acaagtatta caaggatcaa 60  
 aagttcaagc agccagagac atatatcaag tttagtgcga cagtagagga cacagtgggt 120  
 gtggagtaca atatggacga ggtagatgaa aagttttata gagagacact atgcaagtac 180  
 tatcccaaaa agaaaaacaa gtcagatgag aacaatcgaa agtgtactga attggagttt 240  
 gaaacaatct gtgacaagtt ggaaaagacc attgaagcac gacaaccgtt tttgtctatg 300  
 gacccacgca acattctatc gtacgaggag ttgtc 335

<210> 19  
 <211> 326  
 <212> ADN  
 <213> Candida albicans

<400> 19  
 agatatagat aatgtattaa atttagaaga agatcaatat gaattaggat ttaaagaagg 60  
 tcaaatacaa ggaacaaaag atcaatattt agaaggaaaa gaatatggtt atcaaactgg 120  
 atttcaacga tttttaatca ttgggtatat tcaagaatta atgaaatttt gggtatccca 180  
 tatagatcaa tataataact cttcttcact tcggaatcat ttgaataatt tggaagatat 240  
 tatggcacia atttctataa cgaatggaga taaagaagtt gaagattatg aaaaaaatat 300  
 taaaaaggca agaaataaat taagag 326

&lt;210&gt; 20

&lt;211&gt; 374

&lt;212&gt; ADN

&lt;213&gt; Candida albicans

&lt;400&gt; 20

cctcaaatg atttccatgc tactaagaat ttttaagaca ttatttggtg atgataatgg 60

agaatttaac ttaagtgaac tagccgattt aattttacga gaaaattccg tggggacatc 120

aattaaaact gaaggaatgg aaagtgatcc atttgcaatt ttaagtgtac ttaatttaac 180

taataattta aatgtggcgc tgattaaaca attgattgaa tatattttta ataaaaccaa 240

atctaaaact gaattcaata ttattttgaa aaaattgtta accaatcaga acgatactac 300

tagagatagg aaatttaaaa ctggattaat aattagtga agattttata atatgccagt 360

tgaagtgatt ccac

374

&lt;210&gt; 21

&lt;211&gt; 35

&lt;212&gt; ADN

&lt;213&gt; Candida albicans

&lt;400&gt; 21

caatttatcc atgttcgnat ctggaaattg atttc

35

&lt;210&gt; 22

&lt;211&gt; 29

&lt;212&gt; ADN

&lt;213&gt; Candida albicans

&lt;400&gt; 22

ccaaatctca aactctctct aattaaaac

29

<210> 23

<211> 38

<212> ADN

<213> Candida albicans

<400> 23

gttcattgttt ggtgactcag agcgtctcaa ctatattg

38

<210> 24

<211> 33

<212> ADN

<213> Candida albicans

<400> 24

tttgataaac acaggctggt cttaaactctgg ctc

33

<210> 25

<211> 32

<212> ADN

<213> Candida albicans

<400> 25

atctctgata tgagatttgg ctttaaaggc ga

32

<210> 26

<211> 32

<212> ADN

<213> Candida albicans

<400> 26

ggctctttttt ccacagctg cctctgttat tg

32

<210> 27

<211> 20

<212> ADN

<213> Candida albicans

<400> 27

attccacac cggacgcttc

20

<210> 28

<211> 20

<212> ADN

<213> Candida albicans

<400> 28

gacaactcct cgtacgatag

20

<210> 29

<211> 20

<212> ADN

<213> Candida albicans

<400> 29

agataatgta ttaaatttag

20

<210> 30

<211> 20

<212> ADN

<213> Candida albicans

<400> 30

ctcttaattt atttcttgcc

20

<210> 31

<211> 20

<212> ADN

<213> Candida albicans

81

<400> 31

cctcaaattg atttccatgc

20

<210> 32

<211> 20

<212> ADN

<213> Candida albicans

<400> 32

gtggaatcac ttcaactggc

20

## REVENDICATIONS

- 1) Polynucléotides isolés contenant chacun une séquence nucléotidique choisie dans le groupe suivant:
- 5 a) un polynucléotide ayant au moins 50 % ou au moins 60 % et de préférence au moins 70 % d'identité avec un polynucléotide codant pour un polypeptide ayant la même fonction et ayant une séquence en acides aminés homologue d'une séquence choisie parmi SEQ ID N° 2, SEQ ID N° 4, SEQ ID
- 10 N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14
- b) un polynucléotide complémentaire du polynucléotide a)
- c) un polynucléotide comprenant au moins 15 bases consécutives du polynucléotide défini en a) et b).
- 2) Polynucléotides selon la revendication 1 tels que ces
- 15 polynucléotides sont des ADN.
- 3) Polynucléotides selon la revendication 1 tels que ces polynucléotides sont des ARN.
- 4) Polynucléotides tels que définis à la revendication 2 comprenant chacun une séquence de nucléotides choisie parmi
- 20 SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 et SEQ ID N° 13.
- 5) Séquences d'ADN telles que définies aux revendications 1, 2 et 4 caractérisées en ce que ces séquences d'ADN sont celles des gènes codant respectivement pour des protéines de
- 25 *Candida albicans* (ayant les mêmes fonctions que les protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361) et contenant chacune une séquence de nucléotides choisie parmi SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 5, SEQ ID N° 7, SEQ ID N° 9, SEQ ID N° 11 et SEQ ID N° 13.
- 30 6) Séquences d'ADN de gènes selon la revendication 5 codant chacune pour une séquence d'acides aminés choisie parmi SEQ ID N° 2, SEQ ID N° 4, SEQ ID N° 6, SEQ ID N° 8, SEQ ID N° 10, SEQ ID N° 12 et SEQ ID N° 14.
- 7) Séquences d'ADN codant pour les protéines PCaDR472,
- 35 PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 selon les revendications 5 et 6 ainsi que les séquences d'ADN qui hybrident avec celles-ci et/ou présentent des homologies significatives avec ces séquences ou des fragments de celles-

- la cellule hôte est E. coli DH5 alpha ou E. coli XL1-Blue.
- 16) Procédé tel que défini à la revendication 13 dans laquelle la cellule hôte est *Saccharomyces cerevisiae*.
- 17) L'un ou plusieurs des plasmides déposés à la CNCM sous
- 5 les numéros I-2214, I-2215, I-2216, I-2217, I-2211, I-2212 et I-2213.
- 18) Procédé de criblage de produits antifongiques caractérisé en ce qu'il comprend une étape où l'on mesure l'activité de l'une des protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527,
- 10 PCaFL024, PCaNL260, PCaDR361, telles que définies à la revendication 11, en présence de chacun des produits dont on souhaite déterminer les propriétés antifongiques et l'on sélectionne les produits ayant un effet inhibiteur sur cette activité.
- 15 19) Utilisation d'un produit sélectionné par le procédé selon la revendication 18 pour l'obtention d'un agent antifongique.
- 20) Utilisation des gènes de *Candida albicans* ou des protéines codées par ces gènes selon l'une des revendications 5 à 11 pour la sélection de produits ayant des propriétés
- 20 antifongiques selon la revendication 19 comme inhibiteurs des protéines de *Candida albicans* codées par ces gènes.
- 21) Compositions pharmaceutiques renfermant à titre de principe actif au moins un inhibiteur des protéines de *Candida albicans* tel que défini à la revendication 20.
- 25 22) Utilisation des compositions telles que définies à la revendication 21 comme agents antifongiques.
- 23) Méthode d'induction d'une réponse immunologique chez un mammifère comprenant l'inoculation à ce mammifère d'un
- 30 polypeptide tel que défini à la revendication 11 ou un fragment de ce polypeptide ayant la même fonction de façon à produire un anticorps permettant de protéger l'animal contre la maladie.
- 24) Anticorps dirigé contre un polypeptide tel que défini à la revendication 11 ou un fragment de ce polypeptide ayant la
- 35 même fonction.
- 25) Anticorps tel que défini à la revendication 24 dirigé contre l'une quelconque des protéines PCaDR472, PCaDR489, 1PCaDR527, 2PCaDR527, PCaFL024, PCaNL260, PCaDR361 ou un



fragment de ces protéines.

26) Utilisation de l'un quelconque des gènes CaDR472, CaDR489, 1CaDR527, 2CaDR527, CaFL024, CaNL260 et CaDR361 ou de l'une quelconque des protéines codées par ces gènes selon

5 l'une des revendications 5 à 11 pour la préparation de compositions utiles pour le diagnostic ou le traitement de maladies causées par la levure pathogène *Candida albicans*.

27) Kit pour le diagnostic d'infections fongiques comprenant une séquence d'ADN tel que défini à l'une des revendications

10 5 à 10 ou une séquence ayant une fonction similaire ou un fragment fonctionnel de cette séquence, le polypeptide codé par cette séquence ou un fragment polypeptidique ayant la même fonction ou un anticorps dirigé contre un tel polypeptide codé par cette séquence d'ADN ou contre un

15 fragment de ce polypeptide.

Comparaison traduction sonde de CaDR472w x YDR472w :

```

1 .....QFIHVRIWKLIFGKTXIELX 20
      ||| :| :| :|
151 NERLQEKQTESLSNYITKMRRDLKILDILQFIHGTLSYLFNHVSDDL 200

21 NSQDLPMEYMIVENVPLLNRFISIPKEYGDLNCSAFVAGIIEGALDNSGF 70
      | : |||:| | | .| || | ...| || |||. | |. |
201 KSSERDNEYMIVDNFPTLTQF..IPGE..NVSCEYFVCGIIKGFLFNAGF 246

71 NADVTAHTVATDANPLRTVFLIKFDDSVLIRESLRF.. 106
      ||| . . |||:| | || || ||
247 PCGVTahrmpQGGHSQRTVYLIQFDRQVLDREGLRFG* 284

```

FIGURE 1

Comparaison traduction sonde de CaDR489 x YDR489w :

```

1 .....FMFGDSERLNYIVRLYIRTRLSK 23
      | : ||| :::| ||| ||||
101 ISMGFLDMQNASNANPPMPNESKLPLLCMETELERLKFVIRSYIRCRLSK 150

24 LNKFTIFYINESSQNDN.....LLSKEERDYIHKYFQILTQLYNNCF 66
      :.||.: |: : .:.:|      |||:| | : | .| |. |
151 IDKFSL.YLRQLNEDENSLISLTDLLSKDEIKYHDTHTSLIWLKLVNDSIL 199

67 KKLPMQMLTYLDDTSGGQSMIVEPDLDPVFIK..... 98
      | :|: | :.|| | .|| ||| .. |||
200 KYPPEELQAINDTGGSVMIDEPDWNKFVFIHVNGPPDGKWNEDPLLQEN 249

```

FIGURE 2

Comparaison traduction sonde de CaDR527 x YDR527w :

```

1 .....ISDMRFGFKGDLIE 14
                                     |:| | | | |:
251 DKLHEKYFPDLPKEVDKDKWMQPVQOKTDKNYIIEDVSECRFDFNGDLV. 299

15 LAPVGDA PKDSSSDIRTHMGLHHHSETPHMAGYTLGELAH LARSTLAGQR 64
      |   |   | | | | |:| |:| |:| | | | | |
300 .....PPTRQIDSTIHSG LHHHSDSPELAGYTIVELEHLARSTFPSQR 342

65 CLSIQTLGRIFHKLGLHKYSILPNQLNDQSFTDESKLSLDFEDRCGT**T 114
      |:| | | | |:| | | | | | | | | | | | | | | |
343 CIAIQTLGRILYKLGQKSYQVLPEIDADTYKEDGSIS.NVMDKIYSMF. 390

115 NYESLKQ*QRLMEKR..... 130
      :::| ...|
391 .WDLIKDG..KVIESLEISSDEKFTRNLSVRNYAIDALWLWKQGGGDFRT 437

```

FIGURE 3

Comparaison traduction sonde de CaFL024 x YFL024c :

```

1 .....IPTPDASRIWPEAHKYYKDQKFKQPETYIK 30
      ||||| | | :| |.: .|||
101 EVHLHRILQMGSGHTKHKDYIPTPDASMTWNEYDKFYTG.SFQETTSYIK 149
      .....
31 FSATVEDTVGVEYNMDEVDEKIFYRETLCYYPKKKNKSDENNRKCTELEF 80
      ||||| | |||| | | | . | || || ||
150 FSATVEDCCGTNYNMDERDETFLNEQV.....NKGSSD....ILTEDEF 189
      .....
81 ETICDKLEKTIEARQPFLSMDPSNILSYEEL..... 111
      | :| | | ||||| | .||:| |
190 EILCSSFEHAIHERQPFLSMDPESILSFEELKPTLIKSDMADFNLRNQLN 239

```

FIGURE 4

Comparaison traduction sonde de CaNL260c.x YNL260c.:

```

1 .....DIDNVLNLEEDQY 13
      | | | . | | | | |
1 MVRNRFIRKMKKNLFKSNHLSYLKSKWKVKITGQIKMDFDNLNLEEQQY 50

14 ELGFKEGQIQGTKDQYLEGKEYGYQTGFQRFLLIIGYIQELMKFWLSHIDQ 63
   : | | | | : | : | | | : | | | | : | : | : | : |
51 QEGFLEGQENENIKQSFLEGKQYGLQVGFQRFLLGQMEGLCDV....IES 96

64 YN.NSSSLRNHLNLEDIMAQISITNGDKEVEDYEKNIKKARNKLR.... 108
   | . | . | : : : : | : . | | . | : : | : : | : | |
97 YGLHSPTLEKNIHTIRTLMKGLKMNNDDSVMEFERVLIKLNKFKRTILI 146

```

FIGURE 5

Comparaison traduction sonde de CaDR361 x YDR361c :

```

1 .....LKLISMLLRIFKTLFG.DDNGEFNLSEIADLILRENS 36
      : |||   |||  :.  || :|||| |
51 IDFDFFGGNPEVDFHALKNLLR...QLFGPQESTRIQLSSLADLIL..GS 95

37 VGTSIKTEGMESDPFAILSVINLTNNLNVAVIKQLIEYILNKTKSKTEFN 86
      |.||||:| ||||: || :. |      | :|: |
96 PTTTIKTDGKESDPYCFLSFVDFKAN.....HLSDYVKYLQKVDMLRS 138

87 IILKKLLTNQNDTTRDRKFKTGLIISERFINMPVEVIP..... 124
      | :: . |      |::||| |||| ||:|
139 TFFKTMIDSGNK.....NCALVLSERLINMPPEVVPPLYKITLEDVAT 181

```

FIGURE 6